

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 21:01:55 ; Search time 168 Seconds
(without alignments)
2079.198 Million cell updates/sec

Title: US-10-781-564-1
Perfect score: 4392
Sequence: 1 MAGNCGVSIALLGVLLGA.....SSKDTDIRPLSTQPMERAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1667569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4392	100.0	836	5	US-10-781-564-1
2	4385	99.8	836	3	US-09-899-569A-4

ALIGNMENTS

RESULT 1
US-10-781-564-1
; Sequence 1, Application US/10781564
; Publication No. US20040247601A1
; GENERAL INFORMATION:
; APPLICANT: Quigley, James P.
; APPLICANT: Hooper, John D.
; APPLICANT: Teata, Jacqueline E.
; APPLICANT: The Scripps Research Institute
; TITLE OR INVENTION: Methods for Diagnosing Cancer and Decreasing Metastasis by Cancer
; FILE REFERENCE: 1361.0360US1
; CURRENT APPLICATION NUMBER: US/10/781.564
; PRIOR FILING DATE: 2004-02-18
; PRIOR FILING DATE: 2003-02-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 836
; TYPE: PRT

ORGANISM: Homo sapiens
US-10-781-564-1

Query Match 100.0%; Score 4392; DB 5; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGNCGVSIALLGVLLGAARLRGAAFAEIALPRESNITVLIKLGPTTLAKCYIVI	60
DB	1	MAGNCGVSIALLGVLLGAARLRGAAFAEIALPRESNITVLIKLGPTTLAKCYIVI	60
QY	61	SKRHITMISGERIVTFSCQSPENHFVIEIQKIDCMGSPCFGEVQLOPSTSLPT	120
DB	61	SKRHITMISGERIVTFSCQSPENHFVIEIQKIDCMGSPCFGEVQLOPSTSLPT	120
QY	121	LNRTFTMDVKAKHSIGLEQPSIRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN	180
DB	121	LNRTFTMDVKAKHSIGLEQPSIRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN	180
QY	181	GTISRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIESVFEGEGSATMSANY	240
DB	181	GTISRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIESVFEGEGSATMSANY	240
QY	241	PEGFPEDELMTQVVPVPAHLPAVSFLNFNLSNCRKEERYEYIIPGSTTNPVFKLBDK	300
DB	241	PEGFPEDELMTQVVPVPAHLPAVSFLNFNLSNCRKEERYEYIIPGSTTNPVFKLBDK	300
QY	301	QPGMAGNFNLSLQCDQDQSPGILRLQFQVLVQHPONESNKIYVVDLSNERAMSLTIE	360
DB	301	QPGMAGNFNLSLQCDQDQSPGILRLQFQVLVQHPONESNKIYVVDLSNERAMSLTIE	360
QY	361	PRPVKQSKFPPGCFVCIESRTCSNLTITSGSHKISFLCDDTLRLMNVKTIISCTDH	420
DB	361	PRPVKQSKFPPGCFVCIESRTCSNLTITSGSHKISFLCDDTLRLMNVKTIISCTDH	420
QY	421	RYCQRKSYSLQVPSDILHLPEVLHDFSWKLVLPKDRSLVLPVPOKLOQHHEKPCNTSF	480
DB	421	RYCQRKSYSLQVPSDILHLPEVLHDFSWKLVLPKDRSLVLPVPOKLOQHHEKPCNTSF	480
QY	481	SYLVASALPSQDLVFGSFCPEGSIKQIQVKNISVTLRTFAPSFOEASROGLTVSFIPY	540
DB	481	SYLVASALPSQDLVFGSFCPEGSIKQIQVKNISVTLRTFAPSFOEASROGLTVSFIPY	540
QY	541	FKERGVFTVTPDTYSKYTLRTPNWDRLPSLITSVMNISVPRDOVACLTFFKENSQVVCQ	600
DB	541	FKERGVFTVTPDTYSKYTLRTPNWDRLPSLITSVMNISVPRDOVACLTFFKENSQVVCQ	600
QY	601	TGRAFMIIQEOQRTAEERIFSLDEDEVLPKPSFHHSFWNNTSNCSPTSQKQDLFLFSVTLT	660
DB	601	TGRAFMIIQEOQRTAEERIFSLDEDEVLPKPSFHHSFWNNTSNCSPTSQKQDLFLFSVTLT	660
QY	661	PRIVDLTVILIAAGVGVLLSALGLIICVKKKKKTKNGPAVINGNINTEMPQPK	720
DB	661	PRIVDLTVILIAAGVGVLLSALGLIICVKKKKKTKNGPAVINGNINTEMPQPK	720
QY	721	KFOGRKQNDNSHVAVIEDTVMYGHLLQDSGSFLQPEVDVYRPFQGTMGVCPSPPTIC	780
DB	721	KFOGRKQNDNSHVAVIEDTVMYGHLLQDSGSFLQPEVDVYRPFQGTMGVCPSPPTIC	780
QY	781	SRAPTAKLATEEPPRSPSESESPYTFPSHNNNDVSSKQDIDLLSTQEMERAE	836
DB	781	SRAPTAKLATEEPPRSPSESESPYTFPSHNNNDVSSKQDIDLLSTQEMERAE	836

RESULT 2
US-09-899-569A-4
; Sequence 4, Application US/09899569A
; Patent No. US20020142003A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020142003Albert Schweifler
; APPLICANT: Marwa Scherl-Moestaguer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Abseher

; TITLE OF INVENTION: Tumorasoziiertes Antigen (B345)
; FILE REFERENCE: 0652.2280001
; CURRENT APPLICATION NUMBER: US/09/899,569A
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: DE 100 33 080.0
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 101 19 294.0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/297,747
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-899-569A-4

Query Match 99.8%; Score 4385; DB 3; Length 836;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPRGAEAFETALPRESNITVLIKGTPTLAKPCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGAARLPRGAEAFETALPRESNITVLIKGTPTLAKPCYIYI 60
QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDMSGPCPGEVQLQSTSLLEPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDMSGPCPGEVQLQSTSLLEPT 120
QY 121 LNRTIIMVYKAHKSIGLEIQESIPLRQIGPESCPDGVTHSISGRIDATVVRIGTFPCSN 180
DB 121 LNRTIIMVYKAHKSIGLEIQESIPLRQIGPESCPDGVTHSISGRIDATVVRIGTFPCSN 180
QY 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEDEGSATLMSANY 240
DB 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEDEGSATLMSANY 240
QY 241 PEGFPEDEIMTWQFVVPALHRAVSFLENFNLSCERKEERVEYYIPGSTTNPDEVFKLEBDK 300
DB 241 PEGFPEDEIMTWQFVVPALHRAVSFLENFNLSCERKEERVEYYIPGSTTNPDEVFKLEBDK 300
QY 301 OPGNAGNPNLSLOCCDDOASPGILRLOFVLYQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 OPGNAGNPNLSLOCCDDOASPGILRLOFVLYQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSRKFVPGCFVCELESRTCCSNLTLTSGSKHKISFLCDDLTRLMMNVEXTISCTDH 420
DB 361 PRPVQSRKFVPGCFVCELESRTCCSNLTLTSGSKHKISFLCDDLTRLMMNVEXTISCTDH 420
QY 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVPKDRLSLVLPAAQLQOHTHEKPCNTSF 480
DB 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVPKDRLSLVLPAAQLQOHTHEKPCNTSF 480
QY 481 SYLVASAIIPSOPLYGSGFCPGGSIKQIQYKONI SVTLRTFAPSFOEASROGLTVSFIPY 540
DB 481 SYLVASAIIPSOPLYGSGFCPGGSIKQIQYKONI SVTLRTFAPSFOEASROGLTVSFIPY 540
QY 541 FAECEGVFTVPTPKSKVYLRTPNMDRGLPSLTSVSNISVPRDOVACLTFPKEKSGVWCQ 600
DB 541 FAECEGVFTVPTPKSKVYLRTPNMDRGLPSLTSVSNISVPRDOVACLTFPKEKSGVWCQ 600
QY 601 TGRAFMIIIOEQRTAEIIFSLDEDLPKPSFHHHSFWNISNCSPTSGKQDLDFSVTLT 660
DB 601 TGRAFMIIIOEQRTAEIIFSLDEDLPKPSFHHHSFWNISNCSPTSGKQDLDFSVTLT 660
QY 661 PRTVLTVLLIAVGGVULLSALGLITCVKKKKKTKNKPAVGIIYNGNINTEMPROPK 720
DB 661 PRTVLTVLLIAVGGVULLSALGLITCVKKKKKTKNKPAVGIIYNGNINTEMPROPK 720
QY 721 KFOGKRKNDSHVYAVIEDTMYGHLLODSGSGFLQPEVDYRPRFOGTWGVCPSPPTIC 780

DB 721 KFOGKRKNDSHVYAVIEDTMYGHLLODSGSGFLQPEVDYRPRFOGTWGVCPSPPTIC 780
QY 781 SRAPYAKUATEEPPPRSPPESESEBPTFTSHPNNGDVSSKDTDIPLLSTOEPMEPAE 836
DB 781 SRAPYAKUATEEPPPRSPPESESEBPTFTSHPNNGDVSSKDTDIPLLANTOEPMEPAE 836

Search completed: May 4, 2006, 21:05:31
Job time : 168 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 20:53:11 ; Search time 188 Seconds

(without alignments)
1953.833 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4392
Sequence: 1 MAGINCGVSIALGLVLLGA.....SSKDTDIPLSTQEPMEPAE 836

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%
Maximum Match 100%

Listing first 1000 summaries

Database :

A: Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4392	100.0	836	8	ADR47303 Human SIM
2	4392	100.0	836	9	Aes01805 Human CUB
3	4385	99.8	836	5	Abp69553 Human pol
4	4385	99.8	836	5	AAM49641 Human tum

ALIGNMENTS

RESULT 1
ADR47303
ID ADR47303 standard; protein; 836 AA.

AC ADR47303;
DT 18-NOV-2004 (first entry)
XX

DE Human SIMA135 protein.

KM tumour marker protein; glycosylated; non-glycosylated; antibody;
KM cytosolic; cancer; metastasis; SIMA135; human.

OS Homo sapiens.
XX

PN	WO2004074481-A1.
XX	
PD	02-SEP-2004.
XX	
PF	18-FEB-2004; 2004WO-EP001556.
XX	
PR	19-FEB-2003; 2003US-0448828P.
XX	
PA	(NOVS) NOVARTIS AG.
PA	(NOVS) NOVARTIS PHARMA GMBH.
PA	(SCRI) SCRIPPS RES INST.
PI	Quigley JP, Hooper JD, Testa JE;
XX	
DR	WPI; 2004-652958/63.
XX	
PT	New glycoprotein antigen SIMA135, useful for diagnosing and treating
PT	cancer, e.g. prostate, colon, gastric, liver, breast, lung, or kidney
PT	cancer.
XX	
PS	Claim 1; SEQ ID NO 1; S1pp; English.
XX	
CC	The invention relates to a novel tumour marker protein. The protein
CC	comprises a sequence of 836 amino acids, ADR47303, or its fragment or
CC	variant, that is glycosylated or non-glycosylated, where the variant has
CC	an amino acid at position 525 which is Arginine and/or the amino acid at
CC	position 827 is Serine. The invention further comprises: an antibody that
CC	binds specifically to the protein, or its fragment or variant, where the
CC	antibody is not mAb 41-2; a pharmaceutical composition comprising the
CC	antibody and a pharmaceutical carrier; and a kit comprising the antibody
CC	and packaging material. The novel protein, its fragments and variants,
CC	and the antibody have cytostatic activity. The antibody is used in a
CC	diagnostic method of cancer, to determine if a test sample contains
CC	metastatic cells, a therapeutic treatment of the human and animal body,
CC	to determine the metastasis modulating ability of an agent, and to
CC	determine if a candidate agent modulates SIMA135 production by a cell.
CC	The antibody is useful for preparing a medicament to inhibit metastasis
CC	by a cancer cell, e.g. epidermoid carcinoma cell, a fibrosarcoma, or a
CC	prostate, colon, gastric, liver, breast, lung, or kidney rhabdoid cancer
CC	cell, or Hep3 cell, in a mammal. The protein or its fragment or variant,
CC	method, and composition are also useful in diagnosing and treating
CC	cancer. This sequence represents a human SIMA135 protein of the
CC	invention.
XX	
SQ	Sequence 836 AA;
	Query Match 100.0%; Score 4392; DB 8; Length 836;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MAGINCGVSIALGLVLLGAARLRGAEAFIALPRESNITVLIKGTPTLLAKPCYVI 60
DB	1 MAGINCGVSIALGLVLLGAARLRGAEAFIALPRESNITVLIKGTPTLLAKPCYVI 60
QY	61 SKRHITMSTKSGRIVTFSCQSPENHFVLEIKNIDCMGSPCFGVQLQPSISLPT 120
DB	61 SKRHITMSTKSGRIVTFSCQSPENHFVLEIKNIDCMGSPCFGVQLQPSISLPT 120
QY	121 LNRFTIMDVKAHKSIGLELOFSIPRLROIIGGSCPCDGVTHSIGRIATVIRIGTFCSN 180
DB	121 LNRFTIMDVKAHKSIGLELOFSIPRLROIIGGSCPCDGVTHSIGRIATVIRIGTFCSN 180
QY	181 GTVSRIKMGEGVKALHLPWFHPNPNVSGFSIANSSIKRLCIIESVFEGSATLMSNY 240
DB	181 GTVSRIKMGEGVKALHLPWFHPNPNVSGFSIANSSIKRLCIIESVFEGSATLMSNY 240
QY	241 PEGFPEDELMTWQVVPVAHLRASVSPFNENSNCRKEERYEYIIPGSTNPEVFKLEDK 300
DB	241 PEGFPEDELMTWQVVPVAHLRASVSPFNENSNCRKEERYEYIIPGSTNPEVFKLEDK 300
QY	301 QPGNMAGNFNLISLQGCDDAQSFGILRLQFVLVQHQPONESNKIYVVDLSNERAMSLTIE 360
DB	301 QPGNMAGNFNLISLQGCDDAQSFGILRLQFVLVQHQPONESNKIYVVDLSNERAMSLTIE 360

```

QY 361 PRPVKQSKRFVPGCVCLSESTCSNLTLTSGSKHKISFLCDLTRLMMANVEKITSCTDH 420
DB 361 PRPVKQSKRFVPGCVCLSESTCSNLTLTSGSKHKISFLCDLTRLMMANVEKITSCTDH 420
QY 421 RYCKQKSYSLQVPSDILHLPEVLHDFSWKLLVPKDRLSLVLPAPKLOOHTHEKPCNTSF 480
DB 421 RYCKQKSYSLQVPSDILHLPEVLHDFSWKLLVPKDRLSLVLPAPKLOOHTHEKPCNTSF 480
QY 481 SYLVASAIPSQDLVYFGSFCPGGSIKQIOVKONISVTLRTFAPSFRQASRQGLTVSFIPY 540
DB 481 SYLVASAIPSQDLVYFGSFCPGGSIKQIOVKONISVTLRTFAPSFRQASRQGLTVSFIPY 540
QY 541 FKKEGVFVTVPTDTSKVKYLRTPNMDRGPSLTYSVMNISVPRDQVACLTFPKERSGVVQ 600
DB 541 FKKEGVFVTVPTDTSKVKYLRTPNMDRGPSLTYSVMNISVPRDQVACLTFPKERSGVVQ 600
QY 601 TGRAFMIIQEOBRTABEIFSLDEVDLPKPSFHHHSFWNINSCSPSTSGKQDLFLFSVTLT 660
DB 601 TGRAFMIIQEOBRTABEIFSLDEVDLPKPSFHHHSFWNINSCSPSTSGKQDLFLFSVTLT 660
QY 661 PRYVDLTVIILAAVGGVLLSALGLIICVAKKKKKTKNKGPAVGIVNGNINTEMPROPK 720
DB 661 PRYVDLTVIILAAVGGVLLSALGLIICVAKKKKKTKNKGPAVGIVNGNINTEMPROPK 720
QY 721 KFOKGRKNDSHVAVIEDTMVYGHLLDSSGSFLQPEVDTRYPRPGTMGVCPSPPTIC 780
DB 721 KFOKGRKNDSHVAVIEDTMVYGHLLDSSGSFLQPEVDTRYPRPGTMGVCPSPPTIC 780
QY 781 SRAPTAKLATEBPPESESEBPYTFSPHNNGVSSKDTDIPLLSTOEPMEPAE 836
DB 781 SRAPTAKLATEBPPESESEBPYTFSPHNNGVSSKDTDIPLLSTOEPMEPAE 836

```

```

RESULT 2
AEA01805
AEA01805 standard; protein; 836 AA.
AC AEA01805;
XX
XX 14-JUL-2005 (first entry)
XX
XX Human CUB domain containing protein 1 (CDCP1).
XX
XX CUB domain containing protein 1; CDCP1; ovary tumor; cytostatic;
XX endocrine disease; genitourinary disease; gynecology and obstetrics;
XX neoplasm; gene therapy; vaccine; diagnosis; drug screening.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 30..667
XX /note="Predicted extracellular domain"
XX
XX MO2005042102-1.
XX
XX 12-MAY-2005.
XX
XX 22-OCT-2004; 2004MO-GB004502.
XX
XX 22-OCT-2003; 2003GB-00024656.
XX
XX (CLUT ) CELLTech R & D LTD.
XX
XX Burgess NA;
XX
XX MPI; 2005-366521/37.
XX N-PSDB; AEA01806.
XX
XX Use of CDCP1 polypeptide or an agent that interacts with or modulates the
XX expression or activity of CDCP1 polypeptide for diagnosing, preventing or
XX treating ovarian cancer or in screening for agents that may treat or
XX prevent the disease.

```

```

XX
PS Claim 6; SEQ ID NO 1; 48bp; English.
XX
CC The present sequence is the protein sequence of human CUB domain
CC containing protein 1 (CDCP1). The invention is based on the finding that
CC CDCP1 represents a novel therapeutic target for the treatment and/or
CC prophylaxis of ovarian cancer. The invention relates to new uses of CDCP1
CC in the diagnosis, screening, treatment and prophylaxis of ovarian cancer,
CC and provides compositions comprising CDCP1, antibodies that are
CC immunospecific for CDCP1, and agents that interact with or modulate the
CC expression or activity of CDCP1 or which modulate the expression a
CC nucleic acid encoding CDCP1. The CDCP1 polypeptide comprises the present
CC sequence or is a derivative of this sequence. The use of an agent that
CC interacts with or modulates the expression or activity of a CDCP1
CC polypeptide for the manufacture of a medicament for the treatment and/or
CC prophylaxis of ovarian cancer is claimed. The agent may be an antibody or
CC its functionally active fragment, derivative or analog, which may be
CC conjugated to a therapeutic moiety, detectable label, second antibody or
CC fragment, effector or reporter molecule, cytotoxic agent or cytokine. A
CC claimed method for the treatment and/or prophylaxis of ovarian cancer
CC involves administering a composition comprising a CDCP1 polypeptide. The
CC composition may be a vaccine. Also claimed are methods of using CDCP1 to
CC screen for anti-ovarian cancer agents, and a method of screening for
CC and/or diagnosis or prognosis of ovarian cancer in a subject, and/or
CC monitoring the effectiveness of ovarian cancer therapy in a subject by
CC detecting and/or quantifying expression of CDCP1 in a sample obtained
CC from the subject.
XX
SQ Sequence 836 AA;
XX
Query Match 100.0%; Score 4392; DB 9; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAGLNCGVSIALLGVLLIGARLPRGAPAFIALPRESNTVYLKGLPTLLARCYVI 60
DB 1 MAGLNCGVSIALLGVLLIGARLPRGAPAFIALPRESNTVYLKGLPTLLARCYVI 60
QY 61 SKRHITMSTKSGERIVTFSCQSPENHFVEIQKNDICMSGPCPFGEVQLOPSTSLPT 120
DB 61 SKRHITMSTKSGERIVTFSCQSPENHFVEIQKNDICMSGPCPFGEVQLOPSTSLPT 120
QY 121 LNRTFIMDVKAHKSIGLEQPSIRLRQIGBSCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIMDVKAHKSIGLEQPSIRLRQIGBSCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKAHLPMFHPNNVSGFSIANSSIRLCITISVFGESATLMSANY 240
DB 181 GTVSRIKMQEGVKAHLPMFHPNNVSGFSIANSSIRLCITISVFGESATLMSANY 240
QY 241 PEGFPEDELMTMOFVVPARHLRASVFLNPNLSNCRKEERYEYIIPGSTTPEVFKLBDK 300
DB 241 PEGFPEDELMTMOFVVPARHLRASVFLNPNLSNCRKEERYEYIIPGSTTPEVFKLBDK 300
QY 301 QPGNAGNFNLSLQCDQDAQSPGILRLQFOVLVOHPQNESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGNAGNFNLSLQCDQDAQSPGILRLQFOVLVOHPQNESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKQSKRFVPGCVCLSESTCSNLTLTSGSKHKISFLCDLTRLMMANVEKITSCTDH 420
DB 361 PRPVKQSKRFVPGCVCLSESTCSNLTLTSGSKHKISFLCDLTRLMMANVEKITSCTDH 420
QY 421 RYCKQKSYSLQVPSDILHLPEVLHDFSWKLLVPKDRLSLVLPAPKLOOHTHEKPCNTSF 480
DB 421 RYCKQKSYSLQVPSDILHLPEVLHDFSWKLLVPKDRLSLVLPAPKLOOHTHEKPCNTSF 480
QY 481 SYLVASAIPSQDLVYFGSFCPGGSIKQIOVKONISVTLRTFAPSFRQASRQGLTVSFIPY 540
DB 481 SYLVASAIPSQDLVYFGSFCPGGSIKQIOVKONISVTLRTFAPSFRQASRQGLTVSFIPY 540
QY 541 FKKEGVFVTVPTDTSKVKYLRTPNMDRGPSLTYSVMNISVPRDQVACLTFPKERSGVVQ 600
DB 541 FKKEGVFVTVPTDTSKVKYLRTPNMDRGPSLTYSVMNISVPRDQVACLTFPKERSGVVQ 600

```

QY 601 TGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHHSFWNINSGSPSTSGKOLDLFSVTLT 660
 DB 601 TGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHHSFWNINSGSPSTSGKOLDLFSVTLT 660
 QY 661 PRVDLTIVIIAAGGVLLLSALGLIICVKKKKKTKNKPAGVINGNINTEMPROPK 720
 DB 661 PRVDLTIVIIAAGGVLLLSALGLIICVKKKKKTKNKPAGVINGNINTEMPROPK 720
 QY 721 KFOGRKNDSHVYAVIEDTMVYGHLLQDSSGSLQPEVDVTRPQGTMGVCPSPPTIC 780
 DB 721 KFOGRKNDSHVYAVIEDTMVYGHLLQDSSGSLQPEVDVTRPQGTMGVCPSPPTIC 780
 QY 781 SRAPTAKLATEEPPRSPSESESEPTTSHPNNGVSSKOTDIPLLSTOEPMEPAE 836
 DB 781 SRAPTAKLATEEPPRSPSESESEPTTSHPNNGVSSKOTDIPLLSTOEPMEPAE 836

RESULT 3
 ABP69553
 ID ABP69553 standard; protein; 836 AA.
 AC ABP69553;
 DT 20-JAN-2003 (first entry)
 DE Human polypeptide SEQ ID NO 1600.
 XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 KW antiarthritic.
 XX
 OS Homo sapiens.
 XX MO200270539-A2.
 XX 12-SEP-2002.
 PD 05-MAR-2002; 2002WO-US005095.
 PF 05-MAR-2002; 2002WO-US005095.
 XX
 PR 05-MAR-2001; 2001US-00799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX
 DR WPI: 2002-759812/82.
 DR N-PSDB; AB211770.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
 PT or coagulation disorders.
 XX
 PS Claim 9; SEQ ID NO 1600; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (1) comprising a
 CC nucleotide sequence selected from any of 948 sequences (AB21119-
 CC AB212066) or their mature protein coding portion, active domain coding
 CC protein or complementary sequences. The polynucleotides are useful for
 CC identifying expressed genes or for physical mapping of human genome. The
 CC encoded polypeptides (ABP6902-ABP6949) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical
 CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,

CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 CC arthritis, etc. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 836 AA;
 Query Match 99.8%; Score 4385; DB 5; Length 836;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGINCYSIALLVLLIGALRLPRGAEPFIALPRESNTIVILKLGPTTLAKPCYIV 60
 DB 1 MAGINCYSIALLVLLIGALRLPRGAEPFIALPRESNTIVILKLGPTTLAKPCYIV 60
 QY 61 SKRHITMLSIKSGRIYFTSCQSPENHFVIEIKNTIDMSGPCPGFVQLQPTSLPT 120
 DB 61 SKRHITMLSIKSGRIYFTSCQSPENHFVIEIKNTIDMSGPCPGFVQLQPTSLPT 120
 QY 121 LNRFTIMDVKAHKSIGLELQPSIRLRQIGGESCPDVTHSIGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDVKAHKSIGLELQPSIRLRQIGGESCPDVTHSIGRIDATVVRIGTFCSN 180
 QY 181 GTVSRIKMQEGVKALHLPMFHPNVSGFSIANSSIKRLCIISVEGEGSATLMSANY 240
 DB 181 GTVSRIKMQEGVKALHLPMFHPNVSGFSIANSSIKRLCIISVEGEGSATLMSANY 240
 QY 241 PEGPEDELMQGVVVAHILRASVPLNPNLSNERKEERYIIPGTTNPEVFKEDK 300
 DB 241 PEGPEDELMQGVVVAHILRASVPLNPNLSNERKEERYIIPGTTNPEVFKEDK 300
 QY 301 QPGMAGNPNLSLOGCDQDQSPGILRLQFQVILQHPONESNKIYVVDLSNERAMSLTIE 360
 DB 301 QPGMAGNPNLSLOGCDQDQSPGILRLQFQVILQHPONESNKIYVVDLSNERAMSLTIE 360
 QY 361 PRPVKQSRKFPVPGCFVLESRTCSNLTITSGSHKISFLCDDLTRLMMNVEKTICTDH 420
 DB 361 PRPVKQSRKFPVPGCFVLESRTCSNLTITSGSHKISFLCDDLTRLMMNVEKTICTDH 420
 QY 421 RYCQKSYSLQVPSDIIHLPELHDFSMKLLVPDRSLVLPVPOKIQOHTHEPCNTSF 480
 DB 421 RYCQKSYSLQVPSDIIHLPELHDFSMKLLVPDRSLVLPVPOKIQOHTHEPCNTSF 480
 QY 481 SYLVASAIPODLYFGSFCPGSIIKOIQVKNISVTLTRFAPSROEASROGLTVSFIY 540
 DB 481 SYLVASAIPODLYFGSFCPGSIIKOIQVKNISVTLTRFAPSROEASROGLTVSFIY 540
 QY 541 FKEEGVFTVPDYSKYVLTLPNMDRGLPSLTSVSNISVPRDVAICTFFKERSGVYCC 600
 DB 541 FKEEGVFTVPDYSKYVLTLPNMDRGLPSLTSVSNISVPRDVAICTFFKERSGVYCC 600
 QY 601 TGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHHSFWNINSGSPSTSGKOLDLFSVTLT 660
 DB 601 TGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHHSFWNINSGSPSTSGKOLDLFSVTLT 660
 QY 661 PRVDLTIVIIAAGGVLLLSALGLIICVKKKKKTKNKPAGVINGNINTEMPROPK 720
 DB 661 PRVDLTIVIIAAGGVLLLSALGLIICVKKKKKTKNKPAGVINGNINTEMPROPK 720
 QY 721 KFOGRKNDSHVYAVIEDTMVYGHLLQDSSGSLQPEVDVTRPQGTMGVCPSPPTIC 780
 DB 721 KFOGRKNDSHVYAVIEDTMVYGHLLQDSSGSLQPEVDVTRPQGTMGVCPSPPTIC 780
 QY 781 SRAPTAKLATEEPPRSPSESESEPTTSHPNNGVSSKOTDIPLLSTOEPMEPAE 836
 DB 781 SRAPTAKLATEEPPRSPSESESEPTTSHPNNGVSSKOTDIPLLSTOEPMEPAE 836

RESULT 4
 AAM49641
 ID AAM49641 standard; protein; 836 AA.

XX AAM49641;
AC 17-MAY-2002 (first entry)
XX
XX Human tumour-associated antigen B345 protein SEQ ID NO 4.
DE
XX Tumour-associated antigen; human; B345; cytostatic; cell communication;
XX cell interaction; signal transduction; metastasis; cancer; colon;
XX immunotherapy; carcinoma; lung; diagnosis.
XX Homo sapiens.
XX MO200204508-A1.
XX 17-JAN-2002.
XX
XX 05-JUL-2001; 2001WO-EP007705.
XX
XX 07-JUL-2000; 2000DE-01033080.
XX 19-APR-2001; 2001DE-01019294.
XX
XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.
PI Schweifer N, Scherl-Mostaguer M, Sommergruber W, Abseher R;
XX WPI: 2002-171704/22.
XX N-P8DB; ABA959507.
XX
XX New tumor-associated antigen B345, useful for diagnosis and immunotherapy
XX of tumors, also related nucleic acid and antibodies.
XX
XX Claim 1; Page 85-88; 102pp; German.
XX
XX This invention describes a novel tumour-associated antigen, designated
XX B345 which has cytostatic activity. B345 is involved in communication,
XX interaction and/or signal transduction with extracellular components and
XX ligands, especially in the metastatic potential of cancers, particularly
XX of the colon. B345 or its immunogenic fragments, also the DNA that
XX encodes it, are useful for immunotherapy of cancer, particularly
XX carcinoma of lung or colon. Antibodies raised against B345 are useful for
XX treatment and diagnosis of cancers that are associated with B345
XX expression, including their use for targeted delivery of cytotoxic or
XX radioactive agents. Probes derived from B345 can be used to detect tumour
XX -specific mutations in the B345 sequence, and can be used to screen for
XX B345 specific modulators. This sequence represents a human B345 tumour-
XX associated antigen described in the invention
XX
XX Sequence 836 AA;
SQ
Query Match 99.8%; Score 4385; DB 5; Length 836;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGLNCGVSIALLGVLLGAAALPRGAEAFETALPRESNITVLILKGTPTLLAKPCYIVI 60
DB 1 MAGLNCGVSIALLGVLLGAAALPRGAEAFETALPRESNITVLILKGTPTLLAKPCYIVI 60
QY 61 SKRHITMTLSIKSGERIVFTFSCSPENHVFIEIQKNIDMSGPCPPGEVQLQPSLTLPT 120
DB 61 SKRHITMTLSIKSGERIVFTFSCSPENHVFIEIQKNIDMSGPCPPGEVQLQPSLTLPT 120
QY 121 LNRFTIMVYKAKHSIGLEIOPSIPLRQIGPESCPDGWTHSISGRIDATVVRIGTFCNS 180
DB 121 LNRFTIMVYKAKHSIGLEIOPSIPLRQIGPESCPDGWTHSISGRIDATVVRIGTFCNS 180
QY 181 GTVSRIKMOEGVKALHLPMFPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIKMOEGVKALHLPMFPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALRLASVSFLNPNLSNCERKEERVYIIPGTTNPEVFKLEBK 300
DB 241 PEGFPEDELMTQFVVPALRLASVSFLNPNLSNCERKEERVYIIPGTTNPEVFKLEBK 300

QY 301 QPGMAGFNLSLQCCDDAOSPGILRLQFOVLVQHONBESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGFNLSLQCCDDAOSPGILRLQFOVLVQHONBESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKOSRKFVPGCFVCLIESRTCSNLTJSGSKHISFLCDDLTRLMNVEXTISCTDH 420
DB 361 PRPVKOSRKFVPGCFVCLIESRTCSNLTJSGSKHISFLCDDLTRLMNVEXTISCTDH 420
QY 421 RYCCQKSYSLQVPSDILHLPELHDFSWKLLVPPDRLSLVVPAQKQOHTHEKPCNFSF 480
DB 421 RYCCQKSYSLQVPSDILHLPELHDFSWKLLVPPDRLSLVVPAQKQOHTHEKPCNFSF 480
QY 481 SYLVASAIPODILFGSFCPGSIKQIOVKONISVTLRTFAPSROEASROGLTVSFIPY 540
DB 481 SYLVASAIPODILFGSFCPGSIKQIOVKONISVTLRTFAPSROEASROGLTVSFIPY 540
QY 541 FKEGVFTVTPDTKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFEKERSGVYQ 600
DB 541 FKEGVFTVTPDTKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFEKERSGVYQ 600
QY 601 TGRAFMIIQEQRTAEEIFSLDEBVLPRKSFHHHSFWNINSNCSPTSGKQDLLEFSVTLT 660
DB 601 TGRAFMIIQEQRTAEEIFSLDEBVLPRKSFHHHSFWNINSNCSPTSGKQDLLEFSVTLT 660
QY 661 PRTVDLTVILIAAVGGVLLSALGLIICVKKKKKKTKNKGPAYGIYNGNINTEMPROPK 720
DB 661 PRTVDLTVILIAAVGGVLLSALGLIICVKKKKKKTKNKGPAYGIYNGNINTEMPROPK 720
QY 721 KFOGRKNDSDHYAVIEDTMYGHLLODSSGSFLQPEVDTYRPPQGTMGVCPSPPTIC 780
DB 721 KFOGRKNDSDHYAVIEDTMYGHLLODSSGSFLQPEVDTYRPPQGTMGVCPSPPTIC 780
QY 781 SRAPTAKLATEPPRSPSESESPYTFSHNNNDVSKDIDILLSTQEMEPAE 836
DB 781 SRAPTAKLATEPPRSPSESESPYTFSHNNNDVSKDIDILLSTQEMEPAE 836

Search completed: May 4, 2006, 20:56:49
Job time : 190 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 21:02:50 ; Search time 28 Seconds

(without alignments)
1381.930 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4392
Sequence: 1 MAGLNCGVSIALLGVLLGA.....SKKTDIPLSTQEMEPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%
Maximum Match 100%
Listing first 1000 summaries

Database : Published Applications AA New*

- 1: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB.pep1.*
- 2: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /SIDS5/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /SIDS5/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 6: /SIDS5/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 7: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
- 8: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
- 9: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
- 10: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 11: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep1.*
- 12: /SIDS5/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4392	100.0	836	11	US-11-050-857-989
2	4385	99.8	836	11	US-11-050-857-988
3	4059	92.4	770	11	US-11-050-857-252

ALIGNMENTS

RESULT 1
US-11-050-857-989
; Sequence 989, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 989

LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-989

Query Match 100.0%; Score 4392; DB 11; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGLNCGVSIALLGVLLGAARLPRGAEAFIALPRESNITVLKLGTPTLIAKPCYIV	60
DB	1	MAGLNCGVSIALLGVLLGAARLPRGAEAFIALPRESNITVLKLGTPTLIAKPCYIV	60
QY	61	SKRHITMLSTISGERIVTFSCQSPENHFVLEIQNIDCMGSPCFRGVOLOPSTSLPT	120
DB	61	SKRHITMLSTISGERIVTFSCQSPENHFVLEIQNIDCMGSPCFRGVOLOPSTSLPT	120
QY	121	LNRTFIMDKAHSIGLELOFSIPRLRQIGBSCPDGVTSHISGRIDATVIRIGTFCSN	180
DB	121	LNRTFIMDKAHSIGLELOFSIPRLRQIGBSCPDGVTSHISGRIDATVIRIGTFCSN	180
QY	181	GTISRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIRLCIIESVFEGSGATLMSANY	240
DB	181	GTISRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIRLCIIESVFEGSGATLMSANY	240
QY	241	PEGFPEDELMTQVVPVPAHLASVSFLNPNISNCERKEERYEYIPGSTNPEVFKLBDK	300
DB	241	PEGFPEDELMTQVVPVPAHLASVSFLNPNISNCERKEERYEYIPGSTNPEVFKLBDK	300
QY	301	QPNMAGNFNLSLOGCDODDQSPGILRLOFVLVQHPONESNKTYVVDLSNERAMSLITE	360
DB	301	QPNMAGNFNLSLOGCDODDQSPGILRLOFVLVQHPONESNKTYVVDLSNERAMSLITE	360
QY	361	PRPVQSKRKFPGCFVCLESRTCSNLTLTSGSKHKSIFLDDTLRLMANNEXTISCTDH	420
DB	361	PRPVQSKRKFPGCFVCLESRTCSNLTLTSGSKHKSIFLDDTLRLMANNEXTISCTDH	420
QY	421	RYCQRKSYSLQVPSDILHLPEVLHDFSWKLLVPKDRSLVLVPAQKLOQHTHEKPCNTSF	480
DB	421	RYCQRKSYSLQVPSDILHLPEVLHDFSWKLLVPKDRSLVLVPAQKLOQHTHEKPCNTSF	480
QY	481	SYLVASAIPSODLVFGSCPCPGSKIQIOVKONISVTLTFAPSFQESROGLTFSFTPY	540
DB	481	SYLVASAIPSODLVFGSCPCPGSKIQIOVKONISVTLTFAPSFQESROGLTFSFTPY	540
QY	541	FKEEGVFTVPTDTSKYVLRTPNMDRGLPSTSVSWNISVPRDQVACLTFFKERSGVVCO	600
DB	541	FKEEGVFTVPTDTSKYVLRTPNMDRGLPSTSVSWNISVPRDQVACLTFFKERSGVVCO	600
QY	601	TGRAFMIIQEQRTAAEIIIFSLDEDLVLPKPSFHHSFWVNISNCSPFSGKOLDLFSVLT	660
DB	601	TGRAFMIIQEQRTAAEIIIFSLDEDLVLPKPSFHHSFWVNISNCSPFSGKOLDLFSVLT	660
QY	661	PRYVDLTYTILAAVGGVLLLSALGLITCCYKTKKKTKNKPAVGINVGNINTEMPROPK	720
DB	661	PRYVDLTYTILAAVGGVLLLSALGLITCCYKTKKKTKNKPAVGINVGNINTEMPROPK	720
QY	721	KFOKGRKNDSHVAVVIEDTWYGHLODSSGSFLQPEVDTRYRPGTMGVCPSPPTIC	780
DB	721	KFOKGRKNDSHVAVVIEDTWYGHLODSSGSFLQPEVDTRYRPGTMGVCPSPPTIC	780
QY	781	SRAPTAKLATEEPPEPSPSESESEBYTFSHPNNGDVSSKTDIPLISTQEMEPAE 836	
DB	781	SRAPTAKLATEEPPEPSPSESESEBYTFSHPNNGDVSSKTDIPLISTQEMEPAE 836	

RESULT 2
US-11-050-857-988
; Sequence 988, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS

;; TITLE OF INVENTION: THEREOF FOR DIAGNOSIS OF OVARIAN CANCER
;; FILE REFERENCE: 1847.1005
;; CURRENT APPLICATION NUMBER: US/11/050,857
;; CURRENT FILING DATE: 2005-01-27
;; NUMBER OF SEQ ID NOS: 1150
;; SEQ ID NO 988
;; LENGTH: 836
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-11-050-857-988

Query Match 99.8%; Score 4385; DB 11; Length 836;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGGAARLPGAAEAFIALPRESNITVLIKLGPTLLAKPCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGGAARLPGAAEAFIALPRESNITVLIKLGPTLLAKPCYIYI 60
QY 61 SKRHITMISIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGCPCPGGEVQLOPSTSLPT 120
DB 61 SKRHITMISIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGCPCPGGEVQLOPSTSLPT 120
QY 121 LNRTFIMDVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIMDVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSIRIKQOEVKALHLPMFHPRNVS GFSIANRSSIKRLCTIESVEGEGSATLMSANY 240
DB 181 GTVSIRIKQOEVKALHLPMFHPRNVS GFSIANRSSIKRLCTIESVEGEGSATLMSANY 240
QY 241 PEGFEDELMTQFVVPALRASVSFLNFNLSNCRKEERVEYIIPGSTTNPBEVKLEDK 300
DB 241 PEGFEDELMTQFVVPALRASVSFLNFNLSNCRKEERVEYIIPGSTTNPBEVKLEDK 300
QY 301 QPNAGNPNLSIQCCDDOASPGILRQFVLYOHQPNESNKIYVVDLSNERAMSLTIE 360
DB 301 QPNAGNPNLSIQCCDDOASPGILRQFVLYOHQPNESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSRKRFVPGFCVLCESRTCSSNLTLTSGSKHISFLCDDLRLMNVNVEKTIISCTDH 420
DB 361 PRPVQSRKRFVPGFCVLCESRTCSSNLTLTSGSKHISFLCDDLRLMNVNVEKTIISCTDH 420
QY 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPRKORLSLVVPAQKLOQHTHEKPCNTSF 480
DB 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPRKORLSLVVPAQKLOQHTHEKPCNTSF 480
QY 481 SYLVASAI PSODLYFGSFCPGGSIKQIOVKONISVTLRTFAPSROEASRQGLTVSFY 540
DB 481 SYLVASAI PSODLYFGSFCPGGSIKQIOVKONISVTLRTFAPSROEASRQGLTVSFY 540
QY 541 FKEBGEFVTVPDKSKVYLRTNMDRGLPSLTSVSNISVPRDQVACLTFFKERSGVVQ 600
DB 541 FKEBGEFVTVPDKSKVYLRTNMDRGLPSLTSVSNISVPRDQVACLTFFKERSGVVQ 600
QY 601 TGRAMITIOEQRTABEIFSLEDEVLPKPSFHHHSFWNINISVPRDQVACLTFFKERSGVVQ 660
DB 601 TGRAMITIOEQRTABEIFSLEDEVLPKPSFHHHSFWNINISVPRDQVACLTFFKERSGVVQ 660
QY 661 PRTVLDLVLLIAVGGVLLLSALGLITICVKKKKKTKNKGPAVGIVNGNINTEMPOK 720
DB 661 PRTVLDLVLLIAVGGVLLLSALGLITICVKKKKKTKNKGPAVGIVNGNINTEMPOK 720
QY 721 KFOQRKONDSHYAVIEDTMYGHLLODSSGSLQPEVDYTRPFOGTMGVCPSPPTIC 780
DB 721 KFOQRKONDSHYAVIEDTMYGHLLODSSGSLQPEVDYTRPFOGTMGVCPSPPTIC 780
QY 781 SAAPPAKLAATEPRPSPESESEPTTFSHPNNGDVSSKOTDIPLLSTOEPMEPAE 836
DB 781 SAAPPAKLAATEPRPSPESESEPTTFSHPNNGDVSSKOTDIPLLSTOEPMEPAE 836

RESULT 3

US-11-050-857-252
;; Sequence 252, Application US/11050857
;; Publication No. US20060040278A1
;; GENERAL INFORMATION:
;; APPLICANT: Compugen Ltd
;; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
;; FILE REFERENCE: 1847.1005
;; CURRENT APPLICATION NUMBER: US/11/050,857
;; CURRENT FILING DATE: 2005-01-27
;; NUMBER OF SEQ ID NOS: 1150
;; SEQ ID NO 252
;; LENGTH: 770
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-11-050-857-252

Query Match 92.4%; Score 4059; DB 11; Length 770;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 768; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 67 MUISKGRIVFTFSCQSPENHFVIEIQKNIDCMGCPCPGGEVQLOPSTSLPTLNRTFI 126
DB 1 MUISKGRIVFTFSCQSPENHFVIEIQKNIDCMGCPCPGGEVQLOPSTSLPTLNRTFI 60
QY 127 MDVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSNIGVSR 186
DB 127 MDVKAHKSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSNIGVSR 186
QY 187 KMQEVKALHLPMFHPRNVS GFSIANRSSIKRLCTIESVEGEGSATLMSANYEGEPE 246
DB 187 KMQEVKALHLPMFHPRNVS GFSIANRSSIKRLCTIESVEGEGSATLMSANYEGEPE 180
QY 247 DELMTQFVVPALRASVSFLNFNLSNCRKEERVEYIIPGSTTNPBEVKLEDKOPGNA 306
DB 247 DELMTQFVVPALRASVSFLNFNLSNCRKEERVEYIIPGSTTNPBEVKLEDKOPGNA 240
QY 307 GNPNLSIQCCDDOASPGILRQFVLYOHQPNESNKIYVVDLSNERAMSLTIEBRPVKQ 366
DB 307 GNPNLSIQCCDDOASPGILRQFVLYOHQPNESNKIYVVDLSNERAMSLTIEBRPVKQ 300
QY 367 SRKFPVPGFCVLCESRTCSSNLTLTSGSKHISFLCDDLRLMNVNVEKTIISCTDHRVQK 426
DB 367 SRKFPVPGFCVLCESRTCSSNLTLTSGSKHISFLCDDLRLMNVNVEKTIISCTDHRVQK 360
QY 427 SYSLQVPSDILHLPELHDFSWKLLVPRKORLSLVVPAQKLOQHTHEKPCNTSFYLVAS 486
DB 427 SYSLQVPSDILHLPELHDFSWKLLVPRKORLSLVVPAQKLOQHTHEKPCNTSFYLVAS 420
QY 487 AIPSDOLYFGSFCPGGSIKQIOVKONISVTLRTFAPSROEASRQGLTVSFYKREGV 546
DB 487 AIPSDOLYFGSFCPGGSIKQIOVKONISVTLRTFAPSROEASRQGLTVSFYKREGV 480
QY 547 FTVTDDTSKLYLRTNMDRGLPSLTSVSNISVPRDQVACLTFFKERSGVVQOTGRAPM 606
DB 547 FTVTDDTSKLYLRTNMDRGLPSLTSVSNISVPRDQVACLTFFKERSGVVQOTGRAPM 540
QY 607 IIOEQRTABEIFSLEDEVLPKPSFHHHSFWNINISVPRDQVACLTFFKERSGVVQOTGRAPM 666
DB 607 IIOEQRTABEIFSLEDEVLPKPSFHHHSFWNINISVPRDQVACLTFFKERSGVVQOTGRAPM 600
QY 667 TVILIAVGGVLLLSALGLITICVKKKKKTKNKGPAVGIVNGNINTEMPOK 726
DB 667 TVILIAVGGVLLLSALGLITICVKKKKKTKNKGPAVGIVNGNINTEMPOK 660
QY 727 KONDHSYAVIEDTMYGHLLODSSGSLQPEVDYTRPFOGTMGVCPSPPTICRAPTA 786
DB 727 KONDHSYAVIEDTMYGHLLODSSGSLQPEVDYTRPFOGTMGVCPSPPTICRAPTA 720
QY 787 KLAATEPRPSPESESEPTTFSHPNNGDVSSKOTDIPLLSTOEPMEPAE 836
DB 787 KLAATEPRPSPESESEPTTFSHPNNGDVSSKOTDIPLLSTOEPMEPAE 770

Fri May 5 16:14:27 2006

us-10-781-564-1.pct.rapbn

Page 3

Search completed: May 4, 2006, 21:06:06
Job time : 28 secs

21

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 20:57:05 ; Search time 45 Seconds
(without alignments)

1787.494 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4392

Sequence: 1 MAGLNCGVSTALLGVLLGA.....SSKDTDIPPLSTQEPMPAPAE 836

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%

Maximum Match 100%

Listing first 1000 summaries

Database :

PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	ID	Description

No matches found

Search completed: May 4, 2006, 21:01:42
Job time : 45 secs

THIS PAGE LEFT BLANK

XX
AC
AAM49641;

17-MAY-2002 (first entry)

Human tumour-associated antigen B345 protein SEQ ID NO 4.

Tumour-associated antigen; human; B345; cytostatic; cell communication; cell interaction; signal transduction; metastasis; cancer; colon; immunotherapy; carcinoma; lung; diagnosis.

Homo sapiens.

W0200204508-241
JAN-2002
05-JUL-2001; 2001WO-EP007705

07-JUL-2000; 2000DE-01033080
19-APR-2001; 2001DE-01019294

19-APR-2001; 2001DE-01019294

(BOEH) BOEHRINGER INGELHEIM INT GMBH

Schweifer N, Scherl-Mostagier M, Sommergruber W, Abseher R,

WPI; 2002-171704/22.

N-PSDB; ABA99507.

New tumor-associated antigen B345, useful for diagnosis and immunotherapy of tumors, also related nucleic acid and antibodies.

Claim 1; Page 85-88; 102pp; German

This invention describes a novel tumour-associated antigen, designated B345 which has cytotoxic activity. B345 is involved in communication and interaction and/or signal transduction with extracellular components and ligands, especially in the metastatic potential of cancers, particularly of the colon. B345 or its immunogenic fragments, also the DNA that encodes it, are useful for immunotherapy of cancer. Particularly associated with B345 are antibodies raised against B345 are useful for carcinoma of lung or colon. Antibodies raised against B345 are useful for treatment and diagnosis of cancers that are associated with B345 expression, including their use for targeted delivery of cytotoxic or radioactive agents. Probes derived from B345 can be used to screen for -specific mutations in the B345 sequence, and can be used to screen for B345 specific modulators. This sequence represents a human B345 tumour-associated antigen described in the invention

Sequence 836 AA:

Query Match	99.8%	Score 4385;	DB 5;	Length 836;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 834;	Conservative	2;	Mismatches 0;	Indels 0;
			Gaps	0;

QY	1	MAGNCGVSTALGVLILGAARLPGRGAEEIALPRESNTIVILKIGTPIILAKPCLIV	80
QY	1	MAGNCGVSTALGVLILGAARLPGRGAEEIALPRESNTIVILKIGTPIILAKPCLIV	80
Db	1	MAGNCGVSTALGVLILGAARLPGRGAEEIALPRESNTIVILKIGTPIILAKPCLIV	60
QY	61	SKRHITMLSIKSGRIVFTFSCSPENHFVIEIQRKIDCMGSCPPREEOLOPSTSLP	120
QY	61	SKRHITMLSIKSGRIVFTFSCSPENHFVIEIQRKIDCMGSCPPREEOLOPSTSLP	120
Db	61	SKRHITMLSIKSGRIVFTFSCSPENHFVIEIQRKIDCMGSCPPREEOLOPSTSLP	120
QY	121	LNRTPIYVWKAKHSIGLELOPSIRLRLRIGBESCPODVTHISGRIDATVARIIGTFCSN	180
QY	121	LNRTPIYVWKAKHSIGLELOPSIRLRLRIGBESCPODVTHISGRIDATVARIIGTFCSN	180
Db	121	LNRTPIYVWKAKHSIGLELOPSIRLRLRIGBESCPODVTHISGRIDATVARIIGTFCSN	180
QY	181	GTVSRIRKQEGVKALHLPMFHPNVSQGSILANSSIRLCIIISVEGEGSATLMANY	240
QY	181	GTVSRIRKQEGVKALHLPMFHPNVSQGSILANSSIRLCIIISVEGEGSATLMANY	240
Db	181	GTVSRIRKQEGVKALHLPMFHPNVSQGSILANSSIRLCIIISVEGEGSATLMANY	240
QY	241	PGSPDELMTQFVVPALIRASVFLFNFLNLSNCRKEERVEYIIPGSTNPBEVFKLEDK	300
QY	241	PGSPDELMTQFVVPALIRASVFLFNFLNLSNCRKEERVEYIIPGSTNPBEVFKLEDK	300
Db	241	PGSPDELMTQFVVPALIRASVFLFNFLNLSNCRKEERVEYIIPGSTNPBEVFKLEDK	300

Qy	301	QPGMAGNENF,SLQCGDQDAQSPGILRLQOVLVQHQNENSKTYVVDLSNERAMSLTIE	360
Db	301	QPGMAGNENF,SLQCGDQDAQSPGILRLQOVLVQHQNENSKTYVVDLSNERAMSLTIE	360
Qy	361	PRPYKQSRKVPYGGFVLCESRTCSNNLTJTSGSGHKJLSFLCDDLTJRLMWNKJTSCTDH	420
Db	361	PRPYKQSRKVPYGGFVLCESRTCSNNLTJTSGSGHKJLSFLCDDLTJRLMWNKJTSCTDH	420
Qy	421	RYCORKSYLQPSDILHLPYELHDFSKMLVPKDRUSLVLPKOKLOOHTHEKCNISF	480
Db	421	RYCORKSYLQPSDILHLPYELHDFSKMLVPKDRUSLVLPKOKLOOHTHEKCNISF	480
Qy	481	SYLVASAIPODLYVGFSPCGSGSIKOIOVKONISVLTLPASPFOQBSROGLJTSFIY	540
Db	481	SYLVASAIPODLYVGFSPCGSGSIKOIOVKONISVLTLPASPFOQBSROGLJTSFIY	540
Qy	541	PKBEQVFTVTPDTKSKVYLRTPNMDRG,PSLTSVSMNISVPRDQVACLTFPKERSGVQC	600
Db	541	PKBEQVFTVTPDTKSKVYLRTPNMDRG,PSLTSVSMNISVPRDQVACLTFPKERSGVQC	600
Qy	601	TGRAMFIIQEBRTABE,IFSLDEDEVLPKRSFHHHSFWNISNCSPTSGQDOLLFSVTLT	660
Db	601	TGRAMFIIQEBRTABE,IFSLDEDEVLPKRSFHHHSFWNISNCSPTSGQDOLLFSVTLT	660
Qy	720	PRTVDLTVLIIAAGGVLLLSALGLIICCVKKKKKKTKNGA,VGIVYNNININTEMROPK	720
Db	720	PRTVDLTVLIIAAGGVLLLSALGLIICCVKKKKKKTKNGA,VGIVYNNININTEMROPK	720
Qy	721	KFOKGKNDSDHYVANI,EDTMVYGHLLQDSGSLFOPEVDTYARFQGTMGVCPSPPTIC	780
Db	721	KFOKGKNDSDHYVANI,EDTMVYGHLLQDSGSLFOPEVDTYARFQGTMGVCPSPPTIC	780
Qy	781	SRAPTAKLATEBP,PPSPSPESBSPYTFSHPNNGDVASKOTDILPLLSQEPMPAE	836
Db	781	SRAPTAKLATEBP,PPSPSPESBSPYTFSHPNNGDVASKOTDILPLNTQEPMPAE	836

Search completed: May 4, 2006, 20:56:45
Job time : 190 secs

THIS PAGE IS BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 21:01:55 ; Search time 168 Seconds

(without alignments)
2079.198 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4392
Sequence: 1 MAGLNGCVSIALLGVLGGA.....SSKDTIDPILSTQEPMEPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%
Maximum Match 100%

Listing first 1000 summaries

Database : Published Applications_AA_Main:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep: *
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep: *
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep: *
4: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep: *
5: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep: *
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	4392	100.0	836	5 US-10-781-564-1	Sequence 1, Appli
2	4385	99.8	836	3 US-09-899-569A-4	Sequence 4, Appli

SUMMARIES

RESULT 1

US-10-781-564-1
; Sequence 1, Application US/10781564
; Publication No. US20040247601A1
; GENERAL INFORMATION:
; APPLICANT: Quigley, James P.
; APPLICANT: Hooper, John D.
; APPLICANT: Testa, Jacqueline E.
; TITLE OF INVENTION: Methods for Diagnosing Cancer and Decreasing Metastasis by Cancer
; FILE REFERENCE: 1361.036US1
; CURRENT APPLICATION NUMBER: US/10/781,564
; PRIOR APPLICATION NUMBER: 2004-02-18
; PRIOR FILING DATE: 2003-02-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 836
; TYPE: PRT

ORGANISM: Homo sapiens
US-10-781-564-1

Query Match 100.0%; Score 4392; DB 5; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGLNGCVSIALLGVLGGAARLPGAAEFAIALPRESNITVLKLGFTLLAKPCYIV	60
DB	1	MAGLNGCVSIALLGVLGGAARLPGAAEFAIALPRESNITVLKLGFTLLAKPCYIV	60
QY	61	SKRHITMLSIKGGERIVFPFSCQSPENHVFIIQKXIDCMSPCPGEGVQLDPSTLLPT	120
DB	61	SKRHITMLSIKGGERIVFPFSCQSPENHVFIIQKXIDCMSPCPGEGVQLDPSTLLPT	120
QY	121	LNRTFLMDVKAKHSIGLEIQFSIPRLRQIGPESCPDGVTHSISGIDATVVRIGFCSN	180
DB	121	LNRTFLMDVKAKHSIGLEIQFSIPRLRQIGPESCPDGVTHSISGIDATVVRIGFCSN	180
QY	181	GVTSRIKQEGVKALHLPWFPRNVSGFSIANRSIKRLCIIESVFEGEGSATLMSANY	240
DB	181	GVTSRIKQEGVKALHLPWFPRNVSGFSIANRSIKRLCIIESVFEGEGSATLMSANY	240
QY	241	PGFPDELMTWQFVVPALRASVSFLFNLSNCRKERVYIIPGSTTNPVFLEDK	300
DB	241	PGFPDELMTWQFVVPALRASVSFLFNLSNCRKERVYIIPGSTTNPVFLEDK	300
QY	301	OPGNAGNPNLSLOGCDODASPGILRLOFVLVQHPQESNKIYVVDLSNERAMSLTE	360
DB	301	OPGNAGNPNLSLOGCDODASPGILRLOFVLVQHPQESNKIYVVDLSNERAMSLTE	360
QY	361	PRPVQSRKFVGCVCVCSRSRTCSNLTLSGSKHISPLCDLTRLMAVNEKTIISCTDH	420
DB	361	PRPVQSRKFVGCVCVCSRSRTCSNLTLSGSKHISPLCDLTRLMAVNEKTIISCTDH	420
QY	421	RYCQKRSYSLQVPSDILHLVPLHDPMSKLVLPKORLSLVYPAQLOQHTHEKPCNTSF	480
DB	421	RYCQKRSYSLQVPSDILHLVPLHDPMSKLVLPKORLSLVYPAQLOQHTHEKPCNTSF	480
QY	481	SYLVASAIPLSODLYGSPFGGSIKQIOVKONISVTLRFAPSPQEAROGITVSFIY	540
DB	481	SYLVASAIPLSODLYGSPFGGSIKQIOVKONISVTLRFAPSPQEAROGITVSFIY	540
QY	541	FKBEGVFTVPTDKSKVYLRTPMNRGLPSLTSVSMNISVPRDOVACLTFKERSGVQO	600
DB	541	FKBEGVFTVPTDKSKVYLRTPMNRGLPSLTSVSMNISVPRDOVACLTFKERSGVQO	600
QY	601	TGRAMTIOEORTRAEIIPLSDEDLVLPKSPFHHSFWVNISNCSPTSGKQDILFVTLT	660
DB	601	TGRAMTIOEORTRAEIIPLSDEDLVLPKSPFHHSFWVNISNCSPTSGKQDILFVTLT	660
QY	661	PTVDLTVILIAAAGGVLLLSALGLIICVYKSKKKTKNKGAVGIIYNININTEMROK	720
DB	661	PTVDLTVILIAAAGGVLLLSALGLIICVYKSKKKTKNKGAVGIIYNININTEMROK	720
QY	721	KFOGRKONDSHVYVVIDTWYGYGLLQDSGSLQPEVDYTRPPQGTGVCPPSPPTIC	780
DB	721	KFOGRKONDSHVYVVIDTWYGYGLLQDSGSLQPEVDYTRPPQGTGVCPPSPPTIC	780
QY	781	SRAPPAKLATEEPPRSPSESESEPYTFSHPNNGDVSSKDTIDPILSTQEPMEPAE	836
DB	781	SRAPPAKLATEEPPRSPSESESEPYTFSHPNNGDVSSKDTIDPILSTQEPMEPAE	836

RESULT 2

US-09-899-569A-4
; Sequence 4, Application US/09899569A
; Patent No. US20020142003A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020142003Albert Schweitzer
; APPLICANT: Marwa Scherl-Mostageer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Abseher

THIS PAGE IS BLANK

Fri May 5 16:14:27 2006

us-10-781-564-1.pct.rapbm

Page 2

Exhibit B page 2 of 2

```

1  TITLE OF INVENTION: Tumorsozsziertes Antigen (B345)
2  FILE REFERENCE: 0652,2280001
3  CURRENT APPLICATION NUMBER: US/09/899,569A
4  PRIOR FILING DATE: 2001-07-06
5  PRIOR APPLICATION NUMBER: DE 100 33 080.0
6  PRIOR FILING DATE: 2000-07-07
7  PRIOR APPLICATION NUMBER: DE 101 19 294.0
8  PRIOR FILING DATE: 2001-04-19
9  PRIOR APPLICATION NUMBER: US 60/243,158
10 PRIOR FILING DATE: 2000-10-25
11 PRIOR APPLICATION NUMBER: US 60/297,747
12 PRIOR FILING DATE: 2001-06-14
13 NUMBER OF SEQ ID NOS: 40
14 SOFTWARE: PatentIn Ver. 2.1
15 SEQ ID NO 4
16 LENGTH: 836
17 TYPE: PRT
18 ORGANISM: Homo sapiens
19 US-09-899-569A-4

```

Query Match	99.8%	Score 4385	DB 3	Length 836
Best Local Similarity	99.8%	Pred. No. 0		
Matches 834	Conservative 2	Mismatches 0	Indels 0	Gaps 0

QY	1	MAGNINGSV	IALLGVLL	GAARLP	REAE	ETAL	PRE	SNT	IVL	KIG	PTTL	AK	PYI	60
Db	1	MAGNINGSV	IALLGVLL	GAARLP	REAE	ETAL	PRE	SNT	IVL	KIG	PTTL	AK	PYI	60
QY	61	SKRHITM	SIKSGER	IVFTF	SCSP	PNH	VI	EIO	KNI	DC	MSG	PC	PE	120
Db	61	SKRHITM	SIKSGER	IVFTF	SCSP	PNH	VI	EIO	KNI	DC	MSG	PC	PE	120
QY	121	LNRTITW	YKAHKS	I	GLE	I	QFSI	PR	LO	I	Q	PG	SC	180
Db	121	LNRTITW	YKAHKS	I	GLE	I	QFSI	PR	LO	I	Q	PG	SC	180
QY	181	GTVSIIT	KIOEGV	YMLHL	PMF	PR	NV	SGS	I	AN	RS	I	K	240
Db	181	GTVSIIT	KIOEGV	YMLHL	PMF	PR	NV	SGS	I	AN	RS	I	K	240
QY	241	PEGFEDE	IMTWQ	FVVP	PAHL	RAS	V	FL	N	F	N	S	N	300
Db	241	PEGFEDE	IMTWQ	FVVP	PAHL	RAS	V	FL	N	F	N	S	N	300
QY	301	QPGNNA	GPN	SL	Q	C	D	D	A	O	S	P	G	360
Db	301	QPGNNA	GPN	SL	Q	C	D	D	A	O	S	P	G	360
QY	361	PRPVQ	SKRFP	PG	CV	C	LES	R	T	S	S	N	L	420
Db	361	PRPVQ	SKRFP	PG	CV	C	LES	R	T	S	S	N	L	420
QY	421	RYCQR	KSY	SL	Q	P	S	D	I	L	H	L	P	480
Db	421	RYCQR	KSY	SL	Q	P	S	D	I	L	H	L	P	480
QY	541	SYLVASAI	P	S	O	D	L	I	F	G	S	F	I	540
Db	541	SYLVASAI	P	S	O	D	L	I	F	G	S	F	I	540
QY	601	FKERG	V	F	V	P	D	T	S	K	Y	L	R	600
Db	601	FKERG	V	F	V	P	D	T	S	K	Y	L	R	600
QY	661	TGRAFM	I	I	O	E	R	T	A	E	I	F	S	660
Db	661	TGRAFM	I	I	O	E	R	T	A	E	I	F	S	660
QY	721	KFOGR	K	D	N	S	H	V	A	I	E	D	T	720
Db	721	KFOGR	K	D	N	S	H	V	A	I	E	D	T	720

D_b 721 KQKGRKONDSHVAVIEDTWYGHLLDQSSGSLQGEVDVTRFRPQGTMGVCPSPPTIC 780

Q_y 781 SRAPFKLATEPPRSPSESESEBYTFSHPNNDVASKDTDLPLSTOEMPEAE 836

D_b 781 SAAPFAKLAATEPPRSPSESESEBYTFSHNNQDVASKDTDLPLATQEMPEAE 836

Search completed: May 4, 2006, 21:05:31
Job time : 168 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 20:53:55 ; Search time 237 Seconds

(without alignments)
2488.699 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4392
Sequence: 1 MAGLNCGVSTALLGVLLIGA.....SSKDTIPPLSTQEPMEPAE 836

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%

Maximum Match 100%

Listing first 1000 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4392	100.0	836	2 Q9H5V8_HUMAN	Q9H5V8 homo sapien
2	4385	99.8	836	2 Q96QU7_HUMAN	Q96QU7 homo sapien

ALIGNMENTS

RESULT 1

Q9H5V8_HUMAN PRELIMINARY; PRT; 836 AA.

AC Q9H5V8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein FLJ22969 (NCSG135).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
RP Watanabe K., Kumagai A., Itakura S., Yamashita M., Tashiro H., Ota T.,
RA Suzuki Y., Ohbayashi M., Nishii T., Shibahara T., Tanaka T.,
RA Nakamura Y., Iino T., Sugano S., Iino T., Sugano S., Iino T.,
RU Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=22547370; PubMed=12660814; DOI=10.1038/sj.onc.1206220;
RA Hooper J.D., Zijlstra A., Aimes R.T., Liang H., Claassen G.F.,
RA Tarin D., Teeta J.E., Quigley J.P.,

RT "Subtractive immunization using highly metastatic human tumor cells
RT identifies SIMA135/CDP1, a 135 kDa cell surface phosphorylated
RT glycoprotein antigen."
RL Oncogene 22:1783-1794(2003).
DR EMBL; AK026622; BAB15511.1; -; mRNA.
DR EMBL; AF468010; AAC33397.1; -; mRNA.
SQ SEQUENCE 836 AA; 92874 MM; 98980475C35C4C8 CRC64;

Query Match 100.0%; Score 4392; DB 2; Length 836;
Best local similarity 100.0%; Pred. No. 2, 4e-312;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAGLNCGVSTALLGVLLIGAARLRPGAEAFIALPRESNTVLKLGPTLLARPCYVI	60
Db	1	MAGLNCGVSTALLGVLLIGAARLRPGAEAFIALPRESNTVLKLGPTLLARPCYVI	60
Qy	61	SKRHITMSTKSGRIYFTFSCQSPENHFVIEIOKNDMSGPCPFGEVQIQPSTSLPT	120
Db	61	SKRHITMSTKSGRIYFTFSCQSPENHFVIEIOKNDMSGPCPFGEVQIQPSTSLPT	120
Qy	121	LNRTFMDVKAHKSIGLEQFSIPRLRQIGPSSCPDGVTHSIGRIDATVVRIGTFCSN	180
Db	121	LNRTFMDVKAHKSIGLEQFSIPRLRQIGPSSCPDGVTHSIGRIDATVVRIGTFCSN	180
Qy	181	GTYSRIKQSGVKKALHPMHPNNVSGFSIANSSIKRLCTIISVPEGSSATLMSANY	240
Db	181	GTYSRIKQSGVKKALHPMHPNNVSGFSIANSSIKRLCTIISVPEGSSATLMSANY	240
Qy	241	PEGFPEDELMTWQGVFVPAHLRASVFLNPNFNSCERKEERYEYIIPGTTPEVFKLEDK	300
Db	241	PEGFPEDELMTWQGVFVPAHLRASVFLNPNFNSCERKEERYEYIIPGTTPEVFKLEDK	300
Qy	301	QPGMAGNFNLSLQCGDQDQSPGILRLQFVLVQHQNQNSKLYVVDLSNERAMSLTIE	360
Db	301	QPGMAGNFNLSLQCGDQDQSPGILRLQFVLVQHQNQNSKLYVVDLSNERAMSLTIE	360
Qy	361	PRPVKQSRKRVPGFVCLBSRTGSSNLTLSGSGHKSIFLCCDLTRLMANVEKTIISCTDH	420
Db	361	PRPVKQSRKRVPGFVCLBSRTGSSNLTLSGSGHKSIFLCCDLTRLMANVEKTIISCTDH	420
Qy	421	RYCORKSYSLQVPSDILHLPELHDFSMKLLVPKDRSLVLPVPAQKQOHTHEPCNTSF	480
Db	421	RYCORKSYSLQVPSDILHLPELHDFSMKLLVPKDRSLVLPVPAQKQOHTHEPCNTSF	480
Qy	481	SYLVASAIIPQDLYFGSFCPGSFKQIQVQNISVTLRTFAPSFRQBSASRQGLTVSFIPIY	540
Db	481	SYLVASAIIPQDLYFGSFCPGSFKQIQVQNISVTLRTFAPSFRQBSASRQGLTVSFIPIY	540
Qy	541	FKERGVTVPDRTSKYVLRTPWMDRGLPSLTSVSNMISVPRDVAALTFPKERSGVVQC	600
Db	541	FKERGVTVPDRTSKYVLRTPWMDRGLPSLTSVSNMISVPRDVAALTFPKERSGVVQC	600
Qy	601	TGRAFMIIQORPRAEITFSLDEVDLPKPSFHHSFWNVSNCSPGTSKQDLDFSTLT	660
Db	601	TGRAFMIIQORPRAEITFSLDEVDLPKPSFHHSFWNVSNCSPGTSKQDLDFSTLT	660
Qy	661	PRIVDLTVIIIAVGGVLLLSALGLIICVKKKKKTKNGPAVINGNINTEMPROPK	720
Db	661	PRIVDLTVIIIAVGGVLLLSALGLIICVKKKKKTKNGPAVINGNINTEMPROPK	720
Qy	721	KFOKGRKDNDSHYAVIEDTMVGHLLQDSGSGFLQEVDTYRPFQGTMGVCPSPPTIC	780
Db	721	KFOKGRKDNDSHYAVIEDTMVGHLLQDSGSGFLQEVDTYRPFQGTMGVCPSPPTIC	780
Qy	781	SRAPTAKLATEBPSPSPSESEPTFSHPNNDVSKOTDIPLLSTQEPMEPAE 836	
Db	781	SRAPTAKLATEBPSPSPSESEPTFSHPNNDVSKOTDIPLLSTQEPMEPAE 836	

RESULT 2

Q96QU7_HUMAN PRELIMINARY; PRT; 836 AA.

ID Q96QU7_HUMAN
AC Q96QU7;

DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CDB domain containing protein 1.
GN Name=CDCP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
RN NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21359860; PubMed=11466621; DOI=10.1038/sj.onc.1204566;
RA Scherl-Mostagier M., Sommergruber W., Abscher R., Hauptmann R.,
RT Ambros P., Schewelfer N.,
RT "Identification of a novel gene, CDCP1, overexpressed in human
RT colorectal cancer."
RL Oncogene 20:4402-4408(2001).
DR EMBL; AY026461; AAK02058.1; -; mRNA.
DR Ensembl; ENSG00000163814; Homo sapiens.
SQ SEQUENCE 836 AA; 92873 MW; FB4D2DBBD35C519 CRC64;

Query Match 99.8%; Score 4385; DB 2; Length 836;
Best Local Similarity 99.8%; Pred. No. 7.8e-312;
Matches 834; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGINCVSIALIGVLLGAAFLPRGAEAFETALPRESNITVLKLGPTTLAKPCYVI 60
DB 1 MAGINCVSIALIGVLLGAAFLPRGAEAFETALPRESNITVLKLGPTTLAKPCYVI 60
QY 61 SKRHITMLSIKGERIVFTFSCQSPENHFVIEIQKINIDCMGSPCPFGEVQLQPTSLAPT 120
DB 61 SKRHITMLSIKGERIVFTFSCQSPENHFVIEIQKINIDCMGSPCPFGEVQLQPTSLAPT 120
QY 121 LNRFTIWDVKAKHSIGLELQFSIPRLRQIGPGESCPCDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWDVKAKHSIGLELQFSIPRLRQIGPGESCPCDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMOEGVKALHLPMFPRNVSGFSIANRSSIKRLCTIESVFEGSGATLMSANY 240
DB 181 GTVSRIKMOEGVKALHLPMFPRNVSGFSIANRSSIKRLCTIESVFEGSGATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALHRSVFLNPNLSNCRKEERVEYIIGSTTNPEVFKLEBK 300
DB 241 PEGFPEDELMTQFVVPALHRSVFLNPNLSNCRKEERVEYIIGSTTNPEVFKLEBK 300
QY 301 QPNNAGNPNLSLQCGDQDAQSGILRLQFOYLVQHPQNESNKIYVVDLSNERAMSLTIE 360
DB 301 QPNNAGNPNLSLQCGDQDAQSGILRLQFOYLVQHPQNESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSKKFPVPGCFVCLSESTCSNLTLTSGSKHISFLCDLTRLMMNVEKTIISCTDH 420
DB 361 PRPVQSKKFPVPGCFVCLSESTCSNLTLTSGSKHISFLCDLTRLMMNVEKTIISCTDH 420
QY 421 RYCQKSKSYLQVPSDILHLPELHDFSWKLVPKDRLSLVLPKQLQOHTHEKPCNTSF 480
DB 421 RYCQKSKSYLQVPSDILHLPELHDFSWKLVPKDRLSLVLPKQLQOHTHEKPCNTSF 480
QY 481 SYLVASAIPSODLYGSPFCPGSGIKQIQVKQNI SVTLRTFAPSFOEASRQGLTFSFIY 540
DB 481 SYLVASAIPSODLYGSPFCPGSGIKQIQVKQNI SVTLRTFAPSFOEASRQGLTFSFIY 540
QY 541 FKEBGVFTVPTPKSKVLRTPNMDRGLPSLTSVSMNISVPRDQVACLTFEKRSGVVCQ 600
DB 541 FKEBGVFTVPTPKSKVLRTPNMDRGLPSLTSVSMNISVPRDQVACLTFEKRSGVVCQ 600
QY 601 TGRAFMITQORTRAEEIFSLDEDLVPKPSFHHHSFWVNISNCSPTSGKQDLDFSVTLT 660
DB 601 TGRAFMITQORTRAEEIFSLDEDLVPKPSFHHHSFWVNISNCSPTSGKQDLDFSVTLT 660
QY 661 PRTVDLTVILLAAVGGVLLLSALGLIICVKKKKKTKNKPAVGIIYNGNINTEMPROPK 720
DB 661 PRTVDLTVILLAAVGGVLLLSALGLIICVKKKKKTKNKPAVGIIYNGNINTEMPROPK 720

QY 721 KFOKGRKNDSHVYAVIEDTIVYGHLLQDSSGSFLQPEVDTPRPPQGTMGVCPSPSPPTIC 780
DB 721 KFOKGRKNDSHVYAVIEDTIVYGHLLQDSSGSFLQPEVDTPRPPQGTMGVCPSPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPPESESEBPYTFSHPNNGDVSSKDTDIPLLSTQEBMEPAE 836
DB 781 SRAPTAKLATEEPPRSPPESESEBPYTFSHPNNGDVSSKDTDIPLLSTQEBMEPAE 836

Search completed: May 4, 2006, 21:00:51
Job time : 238 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 22:05:42 ; Search time 134 Seconds
(without alignments)
2741.199 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4377

Sequence: 1 MAGNCGSIALGLVLLGA.....SSKTDIPDLXQEMEPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*

1:	geneeqp1980s:*
2:	geneeqp1990s:*
3:	geneeqp2000s:*
4:	geneeqp2001s:*
5:	geneeqp2002s:*
6:	geneeqp2003as:*
7:	geneeqp2003bs:*
8:	geneeqp2004s:*
9:	geneeqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4375	99.9	836	ADR47303	Adt47303 Human STM
2	4375	99.9	836	Aea01805	Aea01805 Human CUB
3	4374	99.9	836	ABP69553	ABP69553 Human POL
4	4374	99.9	836	AAM49641	AAM49641 Human tum
5	3625.5	82.8	709	AAV91456	AAV91456 Human sec
6	3625.5	82.8	709	ADL71528	ADL71528 Human hum
7	3454	78.9	749	AAM49640	AAM49640 Human tum
8	3395	77.6	649	AAB95558	AAB95558 Human pro
9	2155	49.2	414	ADL71669	ADL71669 Novel hum
10	2155	49.2	443	AAV91592	AAV91592 Human sec
11	2155	49.2	443	ADL71667	ADL71667 Novel hum
12	1779	40.6	343	AAU12252	AAU12252 Human PRO
13	1779	40.6	343	ABO17686	ABO17686 Novel hum
14	1779	40.6	343	ABU80950	ABU80950 Human PRO
15	1779	40.6	343	ABU66650	ABU66650 Human PRO
16	1779	40.6	343	ABU59731	ABU59731 Novel sec
17	1779	40.6	343	ABO24921	ABO24921 Human sec
18	1779	40.6	343	ABU66926	ABU66926 Human sec
19	1779	40.6	343	ADA56581	ADA56581 Novel hum
20	1779	40.6	343	ADH76112	ADH76112 Human PRO
21	1779	40.6	343	ADA18762	ADA18762 Human PRO
22	1779	40.6	343	ADA61385	ADA61385 Homo sapi
23	1779	40.6	343	ADB19170	ADB19170 Novel hum
24	1779	40.6	343	ADB27711	ADB27711 Human PRO

25	1779	40.6	343	6	ADA86190	ADA86190 Novel hum
26	1779	40.6	343	6	ADB15754	ADB15754 Human PRO
27	1779	40.6	343	6	ADA47540	ADA47540 Human PRO
28	1779	40.6	343	6	ADA67335	ADA67335 Human PRO
29	1779	40.6	343	6	ADB30342	ADB30342 Human PRO
30	1779	40.6	343	6	ADA85638	ADA85638 Novel hum
31	1779	40.6	343	6	ADA96850	ADA96850 Human PRO
32	1779	40.6	343	6	ADA79154	ADA79154 Human PRO
33	1779	40.6	343	6	ADA87293	ADA87293 Novel hum
34	1779	40.6	343	6	ADB16495	ADB16495 Human PRO
35	1779	40.6	343	6	ADA91587	ADA91587 Novel hum
36	1779	40.6	343	6	ADB14650	ADB14650 Human PRO
37	1779	40.6	343	6	ADB19611	ADB19611 Novel hum
38	1779	40.6	343	6	ADA93826	ADA93826 Human PRO
39	1779	40.6	343	6	ADB19722	ADB19722 Novel hum
40	1779	40.6	343	6	ADB13034	ADB13034 Human PRO
41	1779	40.6	343	6	ABO43229	ABO43229 Novel hum
42	1779	40.6	343	6	ADA74288	ADA74288 Human PRO
43	1779	40.6	343	6	ADB24521	ADB24521 Human PRO
44	1779	40.6	343	6	ADA82045	ADA82045 Human PRO
45	1779	40.6	343	6	ADA75008	ADA75008 Human PRO

ALIGNMENTS

RESULT 1
ADR47303
ID ADR47303 standard; protein; 836 AA.
XX
AC ADR47303;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human SIMA135 protein.
XX
DE
XX
KW tumour marker protein; glycosylated; non-glycosylated; antibody;
KW cytosolic; cancer; metastasis; SIMA135; human.
XX
OS Homo sapiens.
XX
PN WO2004074481-A1.
XX
PD 02-SEP-2004.
XX
PF 18-FEB-2004; 2004WO-EP001556.
XX
PR 19-FEB-2003; 2003US-0448828P.
XX
PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS PHARMA GMBH.
PA (SCRI) SCRIPPS RES INST.
XX
PI Quigley JP, Hooper JD, Testa JE;
XX
DR WPI: 2004-652958/63.
XX
PT New glycoprotein antigen SIMA135, useful for diagnosing and treating
PT cancer, e.g. prostate, colon, gastric, liver, breast, lung, or kidney
PT cancer.
XX
PS Claim 1: SEQ ID NO 1; 51pp; English.
XX
The invention relates to a novel tumour marker protein. The protein
comprises a sequence of 836 amino acids, ADR47303, or its fragment or
variant, that is glycosylated or non-glycosylated, where the variant has
an amino acid at position 525 which is Arginine and/or the amino acid at
position 827 is Serine. The invention further comprises: an antibody that
binds specifically to the protein, or its fragment or variant, where the
antibody is not mAb 41-2; a pharmaceutical composition comprising the
antibody and a pharmaceutical carrier; and a kit comprising the antibody
and packaging material. The novel protein, its fragments and variants,
and the antibody have cytostatic activity. The antibody is used in a

CC diagnostic method of cancer, to determine if a test sample contains
CC metastatic cells, a therapeutic treatment of the human and animal body,
CC to determine the metastasis modulating ability of an agent, and to
CC determine if a candidate agent modulates SIMA13 production by a cell.
CC The antibody is useful for preparing a medicament to inhibit metastasis
CC by a cancer cell, e.g. epidermoid carcinoma cell, a fibrosarcoma, or a
CC prostate, colon, gastric, liver, breast, lung, or kidney rhabdoid cancer
CC cell, or Hep3 cell, in a mammal. The protein or its fragment or variant,
CC method, and composition are also useful in diagnosing and treating
CC cancer. This sequence represents a human SIMA13 protein of the
CC invention.

CC
XX
SQ Sequence 836 AA;

Query Match 99.9%; Score 4375; DB 8; Length 836;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPAGAEAFETALPRESNITVLKLGPTLLAKPCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGAARLPAGAEAFETALPRESNITVLKLGPTLLAKPCYIYI 60
QY 61 SKRHITMLSIKSGERIVFTSCSPENHFVIEIQKNIDMSGPCPEGEVQLQSTSLPT 120
DB 61 SKRHITMLSIKSGERIVFTSCSPENHFVIEIQKNIDMSGPCPEGEVQLQSTSLPT 120
QY 121 LNRITIMVYKAHKSIGLEIFSIPLRQIGPESCPDGYTHISGRIDATVYRIGTFCSN 180
DB 121 LNRITIMVYKAHKSIGLEIFSIPLRQIGPESCPDGYTHISGRIDATVYRIGTFCSN 180
QY 181 GTVSRITKQEGVMALHLPMFAPRVNVSFISINRSIKLCTIESFEGEGSATIMSANY 240
DB 181 GTVSRITKQEGVMALHLPMFAPRVNVSFISINRSIKLCTIESFEGEGSATIMSANY 240
QY 241 PEGPEDEIMTQFVVPALHRSVSFLNFNISNCEKEEREVEYIIPGSTTNEVEFKLEDK 300
DB 241 PEGPEDEIMTQFVVPALHRSVSFLNFNISNCEKEEREVEYIIPGSTTNEVEFKLEDK 300
QY 301 QGNNAGNPNLSLOCDODNOSPGLIRLOFQVLYVQHPOESNKIYVVDLSNERAMSLTE 360
DB 301 QGNNAGNPNLSLOCDODNOSPGLIRLOFQVLYVQHPOESNKIYVVDLSNERAMSLTE 360
QY 361 PRPVQSRKFVPGCEVFCLESRTCSNLTLSGSKHKISFLCDLFTLMANVEKITISCTDH 420
DB 361 PRPVQSRKFVPGCEVFCLESRTCSNLTLSGSKHKISFLCDLFTLMANVEKITISCTDH 420
QY 421 RYCQRKSYSLQVPSDILHLPELHDFSMLVVKDLVLVPAQKLOOHTHEKPCNTSF 480
DB 421 RYCQRKSYSLQVPSDILHLPELHDFSMLVVKDLVLVPAQKLOOHTHEKPCNTSF 480
QY 481 SYLVASATPSODLYFSGPCPGSIKQIYQKQNSVTLRTFAPSFOEASROGLTVSFIPIY 540
DB 481 SYLVASATPSODLYFSGPCPGSIKQIYQKQNSVTLRTFAPSFOEASROGLTVSFIPIY 540
QY 541 FPEBEGFTVPTPKSKVYLRTPNMDGLPSLTSVSNISVPRDQVACLTFPERSGVQO 600
DB 541 FPEBEGFTVPTPKSKVYLRTPNMDGLPSLTSVSNISVPRDQVACLTFPERSGVQO 600
QY 601 TGRAPMTIOEORTRAEFTSLDEDLVLRKPSFHHSFWANI SNCSPTSGKQDLLESVTLT 660
DB 601 TGRAPMTIOEORTRAEFTSLDEDLVLRKPSFHHSFWANI SNCSPTSGKQDLLESVTLT 660
QY 661 PRTVULTVLLIANGGVLLLSALGLITICVKKKKKKTKKGPVGVYNNINTEMROK 720
DB 661 PRTVULTVLLIANGGVLLLSALGLITICVKKKKKKTKKGPVGVYNNINTEMROK 720
QY 721 KFOGRKNDSHVYAVIEDTMVYGHLLQDSGSLQPEVDYTRPFGTGVCPSPSPITIC 780
DB 721 KFOGRKNDSHVYAVIEDTMVYGHLLQDSGSLQPEVDYTRPFGTGVCPSPSPITIC 780
QY 781 SPAPTAKLATEBPPRSPSESESEPTFSHPNNGDVSSKQTDIPLIXTOEPMEPAR 836
DB 781 SPAPTAKLATEBPPRSPSESESEPTFSHPNNGDVSSKQTDIPLIXTOEPMEPAR 836

RESULT 2
AEA01805
ID AEA01805 standard; protein; 836 AA.

AC AEA01805;
XX 14-UTL-2005 (first entry)
DT

XX Human CUB domain containing protein 1 (CDCP1).

XX CUB domain containing protein 1; CDCP1; ovary tumor; cytostatic;
KW endocrine disease; genitourinary disease; gynecology and obstetrics;
KW neoplasm; gene therapy; vaccine; diagnosis; drug screening.

OS Homo sapiens.

Key Location/Qualifiers
FT Domain 30..667
FT /note="Predicted extracellular domain"

PN WO2005042102-A1.

PD 12-MAY-2005.

PP 22-OCT-2004; 2004WO-GB004502.

PR 22-OCT-2003; 2003GB-00024656.

XX (CLLT) CELUTECH R & D LTD.

PI Burgess NA;

DR WPI, 2005-366521/37.

DR N-PSDB; AEA01806.

PT Use of CDCP1 polypeptide or an agent that interacts with or modulates the
PT expression or activity of CDCP1 polypeptide for diagnosing, preventing or
PT treating ovarian cancer or in screening for agents that may treat or
PT prevent the disease.

PS Claim 6; SEQ ID NO 1; 48bp; English.

XX The present sequence is the protein sequence of human CUB domain
CC containing protein 1 (CDCP1). The invention is based on the finding that
CC CDCP1 represents a novel therapeutic target for the treatment and/or
CC prophylaxis of ovarian cancer. The invention relates to new uses of CDCP1
CC in the diagnosis, screening, treatment and prophylaxis of ovarian cancer,
CC and provides compositions comprising CDCP1, antibodies that are
CC immunospecific for CDCP1, and agents that interact with or modulate the
CC expression or activity of CDCP1 or which modulate the expression a
CC nucleic acid encoding CDCP1. The CDCP1 polypeptide comprises the present
CC sequence or is a derivative of this sequence. The use of an agent that
CC interacts with or modulates the expression or activity of a CDCP1
CC polypeptide for the manufacture of a medicament for the treatment and/or
CC prophylaxis of ovarian cancer is claimed. The agent may be an antibody or
CC its functionally active fragment, derivative or analog, which may be
CC conjugated to a therapeutic moiety, detectable label, second antibody or
CC fragment, effector or reporter molecule, cytotoxic agent or cytokine. A
CC claimed method for the treatment and/or prophylaxis of ovarian cancer
CC involves administering a composition comprising a CDCP1 polypeptide. The
CC composition may be a vaccine. Also claimed are methods of using CDCP1 to
CC screen for anti-ovarian cancer agents, and a method of screening for
CC and/or diagnosis or prognosis of ovarian cancer in a subject, and/or
CC monitoring the effectiveness of ovarian cancer therapy in a subject by
CC detecting and/or quantifying expression of CDCP1 in a sample obtained
CC from the subject.

XX
SQ Sequence 836 AA;

Query Match 99.9%; Score 4375; DB 9; Length 836;
Best Local Similarity 99.6%; Pred. No. 0;

Matches	833: Conservative	0: Mismatches	3: Indels	0: Gaps	0:
QY	1	MAGNCGVSIALLGVLLGGAARLRGAEAFELAPRESNTIVLTKLGPTLLAPCYVI	60		
Db	1	MAGNCGVSIALLGVLLGGAARLRGAEAFELAPRESNTIVLTKLGPTLLAPCYVI	60		
QY	61	SKRHITMSTIKSGERIVFTFSCQSPENHFVEIQKNIDCMGSPCPGEGVQLQPSSTLLPT	120		
Db	61	SKRHITMSTIKSGERIVFTFSCQSPENHFVEIQKNIDCMGSPCPGEGVQLQPSSTLLPT	120		
QY	121	LNRTFIWDVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN	180		
Db	121	LNRTFIWDVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN	180		
QY	181	GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIESVEFGESATLMSANY	240		
Db	181	GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIESVEFGESATLMSANY	240		
QY	241	PEGPEDELMTWQFVVPFAHLRASVSFLNPNLSNCRKEERVEYIIPGSTTNPEVFKLBDK	300		
Db	241	PEGPEDELMTWQFVVPFAHLRASVSFLNPNLSNCRKEERVEYIIPGSTTNPEVFKLBDK	300		
QY	301	QPGNAGNPNLSLQSCDDAOSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE	360		
Db	301	QPGNAGNPNLSLQSCDDAOSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE	360		
QY	361	PRPVQSRKRPVPGCVCLIESRTCSNLTITSGSKHKISFLCDLTRLMMANVEKITSCTDH	420		
Db	361	PRPVQSRKRPVPGCVCLIESRTCSNLTITSGSKHKISFLCDLTRLMMANVEKITSCTDH	420		
QY	421	RYCOKRSYLOVPSDILHLPVELHDFSMKLVPRKRLSLVVPACKLOQTHKEKCNISF	480		
Db	421	RYCOKRSYLOVPSDILHLPVELHDFSMKLVPRKRLSLVVPACKLOQTHKEKCNISF	480		
QY	481	SYLVASAIPODLVPGSFCPGSGIKQIVKONISVTLTRFAPSFYQASRQGLTVSFIPY	540		
Db	481	SYLVASAIPODLVPGSFCPGSGIKQIVKONISVTLTRFAPSFYQASRQGLTVSFIPY	540		
QY	541	FKEBGFVTPPTTSKVTLRTPNMDRGLPSLTYSVMNISVPRDQVACLTFEKERSGVVCO	600		
Db	541	FKEBGFVTPPTTSKVTLRTPNMDRGLPSLTYSVMNISVPRDQVACLTFEKERSGVVCO	600		
QY	601	TGRAFMIIQEOBTRAEIIFSLDEVDLPKPSFHHSFWNINSCSTSKOJDLFSVTLT	660		
Db	601	TGRAFMIIQEOBTRAEIIFSLDEVDLPKPSFHHSFWNINSCSTSKOJDLFSVTLT	660		
QY	661	PRIVDLTVILIAAVGGVLLSALGLIICVAKKKKKTKNKGPAGVINYXNINTEMROPK	720		
Db	661	PRIVDLTVILIAAVGGVLLSALGLIICVAKKKKKTKNKGPAGVINYXNINTEMROPK	720		
QY	721	KFOGRKNDSHVAVIEDTMYGHLQDSSGSFLQPEVDYTRPFGTMGVCPPSPPTIC	780		
Db	721	KFOGRKNDSHVAVIEDTMYGHLQDSSGSFLQPEVDYTRPFGTMGVCPPSPPTIC	780		
QY	781	SRAPIAKLATEEPPRSPSESESEPTESHNNGVSSKDDIPILSTQEPMEPAE	836		
Db	781	SRAPIAKLATEEPPRSPSESESEPTESHNNGVSSKDDIPILSTQEPMEPAE	836		

KW	arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW	antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW	haemostatic; vulnereary; fungicide; antibacterial; vitrucide; protozoacide;
OS	antiarthritic.
XX	
XX	Homo sapiens.
XX	
XX	WO2002/0539-A2.
PD	12-SEP-2002.
XX	
XX	05-MAR-2002; 2002WO-US005095.
PF	
XX	
PR	05-MAR-2001; 2001US-00799451.
XX	
XX	(HYSE-) HYSEQ INC.
PA	
PI	Tang YF, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI	Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI	Wehrman T, Wang J, Wang D, Drmanac RT;
XX	
DR	WPI; 2002-759812/82.
DR	N-PSDB; AB211770.
XX	
PT	New polynucleotides comprising sequences assembled from expressed
PT	sequence tags (ESTs), useful for treating cell-proliferative,
PT	neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT	or coagulation disorders.
XX	
XX	Claim 9; SEQ ID NO 1600; 1012pp + Sequence Listing; English.
PS	
CC	The invention relates to an isolated polynucleotide (I) comprising a
CC	nucleotide sequence selected from any of 948 sequences (AB21119-
CC	AB212066) or their mature protein coding portion, active domain coding
CC	protein or complementary sequences. The polynucleotides are useful for
CC	identifying expressed genes or for physical mapping of human genome. The
CC	encoded polypeptides (ABP6902-ABP69849) are useful as molecular weight
CC	markers, as a food supplement, for generating antibodies, in medical
CC	imaging, screening and diagnostic assays and for treating cell-
CC	proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC	or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC	diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC	platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC	or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC	arthritis, etc. Note: The sequence data for this patent did not form part
CC	of the printed specification, but was obtained in electronic format
CC	directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 836 AA:
QY	1 MAGNCGVSIALLGVLLGGAARLRGAEAFELAPRESNTIVLTKLGPTLLAPCYVI
Db	1 MAGNCGVSIALLGVLLGGAARLRGAEAFELAPRESNTIVLTKLGPTLLAPCYVI
QY	61 SKRHITMSTIKSGERIVFTFSCQSPENHFVEIQKNIDCMGSPCPGEGVQLQPSSTLLPT
Db	61 SKRHITMSTIKSGERIVFTFSCQSPENHFVEIQKNIDCMGSPCPGEGVQLQPSSTLLPT
QY	121 LNRTFIWDVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN
Db	121 LNRTFIWDVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN
QY	181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIESVEFGESATLMSANY
Db	181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIESVEFGESATLMSANY
QY	241 PEGPEDELMTWQFVVPFAHLRASVSFLNPNLSNCRKEERVEYIIPGSTTNPEVFKLBDK
Db	241 PEGPEDELMTWQFVVPFAHLRASVSFLNPNLSNCRKEERVEYIIPGSTTNPEVFKLBDK

Query Match 99.9%; Score 4374; DB 5; Length 836;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 301 OPGNAGNFNLSTLQCCDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 OPGNAGNFNLSTLQCCDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSRKFVPGCFVCLSESRCCSNLTLTSGSKHISFLCDLTRLMMANVEKITSCTDH 420
DB 361 PRPVQSRKFVPGCFVCLSESRCCSNLTLTSGSKHISFLCDLTRLMMANVEKITSCTDH 420
QY 421 RYCQKRSYSLQVPSDILHLPELHDFSMWKLVPKORLSLVLPKQKLOQTHEKPCNTSF 480
DB 421 RYCQKRSYSLQVPSDILHLPELHDFSMWKLVPKORLSLVLPKQKLOQTHEKPCNTSF 480
QY 481 SYLVASAIPSODLYPGSFPCGSGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFLPY 540
DB 481 SYLVASAIPSODLYPGSFPCGSGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFLPY 540

QY 541 FKEBGFVTVPDTKSKVYLRTPNMWRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
DB 541 FKEBGFVTVPDTKSKVYLRTPNMWRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
QY 601 TGRAFMIIOBORTRAEELFSLDEVDLPRPSFHHSFWNISNCSPTSGKQDLPLFSVLT 660
DB 601 TGRAFMIIOBORTRAEELFSLDEVDLPRPSFHHSFWNISNCSPTSGKQDLPLFSVLT 660
QY 661 PRTVDLTVLLIAVGGVLLLSALGLIICVKKKKKTKNGPANGIYNNXNITEMPROPK 720
DB 661 PRTVDLTVLLIAVGGVLLLSALGLIICVKKKKKTKNGPANGIYNNXNITEMPROPK 720
QY 721 KFQGRKNDSHVAVIEDTMYVGHLLQDSSGSLQPEVDYTRPFGTGWGCPSPPTIC 780
DB 721 KFQGRKNDSHVAVIEDTMYVGHLLQDSSGSLQPEVDYTRPFGTGWGCPSPPTIC 780
QY 781 SRAPTAKLATEBPSPSPSESESEPTFSHPNNGDVSSKOTDIPILXTOEPMEPAE 836
DB 781 SRAPTAKLATEBPSPSPSESESEPTFSHPNNGDVSSKOTDIPILXTOEPMEPAE 836

RESULT 4

AAM49641
ID AAM49641 standard; protein; 836 AA.

AC AAM49641;
XX
DT 17-MAY-2002 (first entry)

DE Human tumour-associated antigen B345 protein SEQ ID NO 4.

KW Tumour-associated antigen; human; B345; cytostatic; cell communication;
KW cell interaction; signal transduction; metastasis; cancer; colon;
KW immunotherapy; carcinoma; lung; diagnosis.

OS Homo sapiens.

PN WO200204508-A1.

PD 17-JAN-2002.

PF 05-JUL-2001; 2001WO-EP007705.

PR 07-JUL-2000; 2000DE-01033080.

PR 19-APR-2001; 2001DE-01019294.

PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.

PI Schweiler N, Schierl-Mostageer M, Sommergruber W, Abseher R;

DR WPI: 2002-171704/22.

DR N-PSDB; ABA959507.

PT New tumor-associated antigen B345, useful for diagnosis and immunotherapy
PT of tumors, also related nucleic acid and antibodies.

PS Claim 1; Page 85-88; 102pp; German.

CC This invention describes a novel tumour-associated antigen, designated
CC B345 which has cytostatic activity. B345 is involved in communication,
CC interaction and/or signal transduction with extracellular components and
CC ligands, especially in the metastatic potential of cancers, particularly
CC of the colon. B345 or its immunogenic fragments, also the DNA that
CC encodes it, are useful for immunotherapy of cancer, particularly
CC carcinoma of lung or colon. Antibodies raised against B345 are useful for
CC treatment and diagnosis of cancers that are associated with B345.
CC expression, including their use for targeted delivery of cytotoxic or
CC radioactive agents. Probes derived from B345 can be used to detect tumour
CC -specific mutations in the B345 sequence, and can be used to screen for
CC B345 specific modulators. This sequence represents a human B345 tumour-
CC associated antigen described in the invention

XX SQ Sequence 836 AA;

Query Match 99.9%; Score 4374; DB 5; Length 836;

Beet Local Similarity 99.6%; Pred. No. 0;

Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAAALPRGAEAFIALPRESNITVLIKLGTPTLLAKPCYIVI 60
DB 1 MAGLNCGVSIALLGVLLGAAALPRGAEAFIALPRESNITVLIKLGTPTLLAKPCYIVI 60
QY 61 SKRHITMSTSGEIVTFSCQSPENHFVIEIQKNDICMGSPCFGEVOLQPSLSLPT 120
DB 61 SKRHITMSTSGEIVTFSCQSPENHFVIEIQKNDICMGSPCFGEVOLQPSLSLPT 120
QY 121 INRFTIMVKAHKSIGLELQFSIPRLRQIGGESCPDVTSHISGRIDATVRLGTFCSN 180
DB 121 INRFTIMVKAHKSIGLELQFSIPRLRQIGGESCPDVTSHISGRIDATVRLGTFCSN 180
QY 181 GTVSRIKMOEGVKALHLPMFHPRNVSQPSIANNSSIRLCIIEBVFEGEGATLMSANY 240
DB 181 GTVSRIKMOEGVKALHLPMFHPRNVSQPSIANNSSIRLCIIEBVFEGEGATLMSANY 240
QY 241 PEGFPEDELMTQFVVPAPHLRASVSFLNPNINSNCRKEERYEYVPGSTNPEVKLEBK 300
DB 241 PEGFPEDELMTQFVVPAPHLRASVSFLNPNINSNCRKEERYEYVPGSTNPEVKLEBK 300
QY 301 OPGNAGNFNLSTLQCCDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 OPGNAGNFNLSTLQCCDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSRKFVPGCFVCLSESRCCSNLTLTSGSKHISFLCDLTRLMMANVEKITSCTDH 420
DB 361 PRPVQSRKFVPGCFVCLSESRCCSNLTLTSGSKHISFLCDLTRLMMANVEKITSCTDH 420
QY 421 RYCQKRSYSLQVPSDILHLPELHDFSMWKLVPKORLSLVLPKQKLOQTHEKPCNTSF 480
DB 421 RYCQKRSYSLQVPSDILHLPELHDFSMWKLVPKORLSLVLPKQKLOQTHEKPCNTSF 480
QY 481 SYLVASAIPSODLYPGSFPCGSGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFLPY 540
DB 481 SYLVASAIPSODLYPGSFPCGSGSIKOIOVKONISVTLRTFAPSFOEASROGLTVSFLPY 540
QY 541 FKEBGFVTVPDTKSKVYLRTPNMWRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
DB 541 FKEBGFVTVPDTKSKVYLRTPNMWRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
QY 601 TGRAFMIIOBORTRAEELFSLDEVDLPRPSFHHSFWNISNCSPTSGKQDLPLFSVLT 660
DB 601 TGRAFMIIOBORTRAEELFSLDEVDLPRPSFHHSFWNISNCSPTSGKQDLPLFSVLT 660
QY 661 PRTVDLTVLLIAVGGVLLLSALGLIICVKKKKKTKNGPANGIYNNXNITEMPROPK 720
DB 661 PRTVDLTVLLIAVGGVLLLSALGLIICVKKKKKTKNGPANGIYNNXNITEMPROPK 720
QY 721 KFQGRKNDSHVAVIEDTMYVGHLLQDSSGSLQPEVDYTRPFGTGWGCPSPPTIC 780
DB 721 KFQGRKNDSHVAVIEDTMYVGHLLQDSSGSLQPEVDYTRPFGTGWGCPSPPTIC 780

QY 761 SRAPTAKLATEPPPPSPSESEBPYTSHPNNGVSSKDTPIPLIXTOEPMEPAE 836
 DB 761 SRAPTAKLATEPPPPSPSESEBPYTSHPNNGVSSKDTPIPLIXTOEPMEPAE 836

RESULT 5
 ID AAY91456 standard; protein; 709 AA.
 XX AAY91456;
 AC AAY91456;
 DT 29-JUN-2000 (first entry)
 XX

Human secreted protein sequence encoded by gene 6 SEQ ID NO:129.

DE Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW anti-HIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
 KW osteopathic; antidiabetic; antibacterial; antidiabetic; antiallergic;
 KW antipneumonia; cardiac; gene therapy; cancer; neurological disorder;
 KW immune disease; inflammation; blood disorder; tumour.

XX Homo sapiens.
 OS
 PN WO200006698-A1.
 XX
 PD 10-FEB-2000.
 XX
 PF 29-JUL-1999; 99WO-US017130.
 XX
 PR 30-JUL-1998; 98US-0094657P.
 PR 05-AUG-1998; 98US-0095486P.
 PR 06-AUG-1998; 98US-0095454P.
 PR 06-AUG-1998; 98US-0095455P.
 PR 12-AUG-1998; 98US-0096319P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.

P1 Komatseulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
 P1 Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
 P1 Soppet DR, Endress GA, Edner R, Olsen HS, Mucenki M;
 XX
 DR MPI; 2000-195282/17.
 DR N-PSDB; AAA26351.

PT New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders.

XX Claim 11; Page 456-459; 634pp; English.

XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
 CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
 CC proteins can have activities based on the tissues and cells they are
 CC expressed in. Examples of the activities are: cytostatic;
 CC immunosuppressive; anti-HIV; antiinflammatory; nootropic; neuroprotective;
 CC antiallergic; osteopathic; antidiabetic; antibacterial; antidiabetic;
 CC antiashma; antipneumonia; and cardiac. The polynucleotides and their
 CC corresponding secreted proteins are useful for preventing, treating or
 CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
 CC pathological conditions can be diagnosed by determining the amount of the
 CC proteins in a sample or by determining the presence of mutations in the
 CC polynucleotides. Specific uses are described for each of the
 CC polynucleotides, based on which tissues they are most highly expressed
 CC in, and include developing products for the diagnosis or treatment of
 CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
 CC Alzheimer's and behavioural disorders, schizophrenia, osteoporosis,
 CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
 CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
 CC reproductive disorders, gastrointestinal disorders, respiratory disorders
 CC and metabolic disorders. The proteins or polynucleotides can also be used

CC as food additives or preservatives. The proteins are also useful for
 CC identifying their binding partners. AAA26337 to AAA26345 and AAY91450 are
 CC sequences used in the exemplification of the present invention

XX Sequence 709 AA;
 SQ

Query Match 82.8%; Score 3625.5; DB 3; Length 709;
 Best Local Similarity 98.3%; Pred. No. 1.9e-308;
 Matches 698; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 1 MAGLNCGVSIALLGVLLGAAALPRGAFAFIALPRESNTITVLKGTPTLLAKPCYVI 60
 DB 1 MAGLNCGVSIALLGVLLGAAALPRGAFAFIALPRESNTITVLKGTPTLLAKPCYVI 60

QY 61 SKRHITMLSIKGERIVTFSCQSPENHFVIEIOKNIDCMGSPCFGEVOLQPTSLPT 120
 DB 61 SKRHITMLSIKGERIVTFSCQSPENHFVIEIOKNIDCMGSPCFGEVOLQPTSLPT 120

QY 121 LNRPTFMVVKAKHKSIGLEQFSIPRLQIGGSCPDQVTHSISGRIDATVVRIGTFCSN 180
 DB 121 LNRPTFMVVKAKHKSIGLEQFSIPRLQIGGSCPDQVTHSISGRIDATVVRIGTFCSN 180

QY 181 GTVSRIRKMQEGVKALHLPMPHPRNVSGFSIANKSSIKRLCTIESVFEGEGSATLMSANY 240
 DB 181 GTVSRIRKMQEGVKALHLPMPHPRNVSGFSIANKSSIKRLCTIESVFEGEGSATLMSANY 240

QY 241 PEGFPEDELMTWQFVVPALHRAVSFLNFNLSNCRKEERVEYYIPGSTNPEVFKLEDK 300
 DB 241 PEGFPEDELMTWQFVVPALHRAVSFLNFNLSNCRKEERVEYYIPGSTNPEVFKLEDK 300

QY 301 OPGMMAGNFNLSLGGCQDDAOSPGLILQPOVLYOHPONEBNKLYVVDLSNERAMSLTIE 360
 DB 301 OPGMMAGNFNLSLGGCQDDAOSPGLILQPOVLYOHPONEBNKLYVVDLSNERAMSLTIE 360

QY 361 PRPVKQSRKFPVPGCFVCLSESRCSNLTLTSGSKHKISFLCDDTLRLMNVKXTISCTDH 420
 DB 361 PRPVKQSRKFPVPGCFVCLSESRCSNLTLTSGSKHKISFLCDDTLRLMNVKXTISCTDH 420

QY 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPKDRLSLVLPKQLQOHTHEKPCNTSF 480
 DB 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPKDRLSLVLPKQLQOHTHEKPCNTSF 480

QY 481 SYLVASAIPODYLFGSPCPGSGIKOIQVKNISIVTLRTFAPSFOEASRGLTVSFTPY 540
 DB 481 SYLVASAIPODYLFGSPCPGSGIKOIQVKNISIVTLRTFAPSFOEASRGLTVSFTPY 540

QY 541 FKESGVFTVPTDTSKYVLRTPNMDRGLPSLTSVSWNISVRRDQVACLTFFKERSGVVCO 600
 DB 541 FKESGVFTVPTDTSKYVLRTPNMDRGLPSLTSVSWNISVRRDQVACLTFFKERSGVVCO 600

QY 601 TGRAFMIIQEQRTAEIIFSLDEEDVLPKPSFHHHSFWVNISNCSPTSGKQDLDFSVTLT 660
 DB 601 TGRAFMIIQEQRTAEIIFSLDEEDVLPKPSFHHHSFWVNISNCSPTSGKQDLDFSVTLT 660

QY 661 PRYDVLTVILIAAVGGVLLSALGLITCCVKKKKKTKNKPNANGVYXN 710
 DB 661 PRYDVLTVILIAAVGGVLLSALGLITCCVKKKKKTKNKPNANGVYXN 710

RESULT 6
 ID ADL71528 standard; protein; 709 AA.
 XX ADL71528;
 AC ADL71528;
 DT 20-MAY-2004 (first entry)
 XX

Novel human secreted protein seqid 132.

DE antiinflammatory; neuroprotective; nootropic; antiparkinsonian;
 KW anticonvulsant; antidiabetic; CNS; gynaecological; antirheumatic;
 KW antiashmatic; anti-HIV; virostatic; endocrine; cytostatic;
 KW immunosuppressive; antiallergic; cardiovascular; respiratory;

KW dermatological; antimicrobial; gastrointestinal; gene therapy;
 KW neurodegenerative disease; behavioral disorder; inflammatory condition;
 KW hyperproliferative disorder; Alzheimer's disease; Parkinson's disease;
 KW Huntington's disease; metabolic disorder; Tay-Sachs disease;
 KW Leash-Nyhan syndrome; reproductive disorder; immunological disorder;
 KW arthritis; asthma; AIDS; endocrine disorder; immune disorder;
 KW Hodgkin's lymphoma; haematopoietic disorder; muscular disorder;
 KW leukaemia; autoimmune disorder; allergy; cancer; cardiovascular disorder;
 KW respiratory disorder; pulmonary disorder; connective tissue disorder;
 KW skin disorder; CNS disorder; congenital disorder; infectious disorder;
 KW gastrointestinal disorder; human; secreted protein.
 XX Homo sapiens.
 XX OS
 XX US2004034196-A1.
 XX PN
 XX 19-FEB-2004.
 XX PD
 XX 27-JAN-2003; 2003US-00351334.
 XX PF
 XX 30-JUL-1998; 98US-0094657P.
 XX PR 05-AUG-1998; 98US-0095486P.
 XX PR 06-AUG-1998; 98US-0095454P.
 XX PR 06-AUG-1998; 98US-0095455P.
 XX PR 12-AUG-1998; 98US-0096319P.
 XX PR 29-JUL-1999; 99MO-US017130.
 XX PR 24-JAN-2000; 2000US-00489847.
 XX PR 25-JAN-2002; 2002US-0350898P.
 XX PA (KOMA/) KOMATSOUIS G A.
 XX PA (ROSE/) ROSEN C A.
 XX PA (RUBE/) RUBEN S M.
 XX PA (DUAN/) DUAN D R.
 XX PA (MOOR/) MOORE P A.
 XX PA (SHIY/) SHI Y.
 XX PA (LAFU/) LAFLEUR D W.
 XX PA (WEIY/) WEI Y.
 XX PI Komatsoulis GA, Rosen CA, Ruben SM, Duan DR, Moore PA, Shi Y;
 PI Lafleur DW, Wei Y;
 DR WPI; 2004-180094/17.
 DR N-PSDB; ADL71412.
 XX PT New human secreted nucleic acid, useful for diagnosing and treating
 PT neurodegenerative, inflammatory, hyperproliferative, metabolic,
 PT reproductive, cardiovascular, respiratory or immunological disorders or
 PT diseases.
 XX PS
 XX Claim 11; SEQ ID NO 132; 234bp; English.
 CC The invention describes an isolated human nucleic acid molecule (1)
 CC comprising a polynucleotide having a nucleotide sequence at least 95%
 CC identical to a sequence polynucleotide fragment of SEQ ID NO: X or of
 CC the cDNA sequence included in ATCC Deposit No: Z, which is hybridisable
 CC to SEQ ID NO: X; or a sequence encoding a polypeptide fragment, domain or
 CC epitope of SEQ ID NO: Y or a polypeptide sequence encoded by the cDNA
 CC sequence included in ATCC Deposit No: Z, which is hybridisable to SEQ ID
 CC NO: X, having a biological activity. The nucleic acids and polypeptides,
 CC pharmaceutical formulations and kits are useful in diagnosing and
 CC treating neurodegenerative diseases states, behavioral disorders,
 CC inflammatory conditions, hyperproliferative disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease or Huntington's diseases), metabolic
 CC disorders (e.g. Tay-Sachs disease or Leash-Nyhan syndrome), reproductive
 CC disorders, immunological disorders (e.g. arthritis, asthma or AIDS),
 CC endocrine and immune disorders (e.g. Hodgkin's lymphoma), haematopoietic
 CC or muscular disorders (e.g. leukaemia), autoimmune disorders, allergy,
 CC cancer, cardiovascular, respiratory or pulmonary disorders, disorders or
 CC conditions afflicting connective tissue, skin disorders, CNS disorders,
 CC congenital disorders, infectious disorders and gastrointestinal
 CC disorders. This is the amino acid sequence of a novel human secreted
 CC protein of the invention. Note: This sequence does not appear in the
 CC printed specification but is available in electronic format from the US

CC patent office at ftp.segdata.uspc.gov/segdata.html?docID=20040034196.
 XX
 XX Sequence 709 AA;
 XX
 XX Query Match 82.8%; Score 3625.5; DB 8; Length 709;
 XX Best Local Similarity 98.3%; Pred. No. 1.9e-308;
 XX Matches 698; Conservative 1; Mismatches 10; Indels 1; Gaps 1;
 QY 1 MAGNCGVSIALLGVLLGAAALPGAAFEIALPRESNITVLILGPTLLAKPCYVI 60
 DB 1 MAGNCGVSIALLGVLLGAAALPGAAFEIALPRESNITVLILGPTLLAKPCYVI 60
 QY 61 SKRHITMISIGSERIVTFSCQSPENHFVIEIQNDICMGSPCFGEVOLPSTSLPT 120
 DB 61 SKRHITMISIGSERIVTFSCQSPENHFVIEIQNDICMGSPCFGEVOLPSTSLPT 120
 QY 121 LNRFTIMDKAKHSIGLELQFSIPRLRQIGGSCPDGVTHSISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDKAKHSIGLELQFSIPRLRQIGGSCPDGVTHSISGRIDATVVRIGTFCSN 180
 QY 181 GTVSRITKQEGVKALHPMFHPRVNGFSIANSSTRLCITBSVFEQGSATLMSANY 240
 DB 181 GTVSRITKMXEGVKALHPMFHPRVNGFSIANSSTRLCITBSVFEQGSATLMSANY 240
 QY 241 PEGFPEDELMTQFVVPALBASVSFLNPNISNCKEKEVEYYIPGSTTPEVFKLBDK 300
 DB 241 PEGFPEDELMTQFVVPALBASVSFLNPNISNCKEKEVEYYIPGSTTPEVFKLBDK 300
 QY 301 QPGNMGFNLSLQCGDDAOSPGLRLQFQVYLVOHPONESNKTIYVDLSNERAMSLTIE 360
 DB 301 QPGNMGFNLSLQCGDDAOSPGLRLQFQVYLVOHPONESNKTIYVDLSNERAMSLTIE 360
 QY 361 PRPVQSKRFVPGCVCSERTCSNLTLTSGSKHKSFLDDDLTRLMNTEKITSCTDH 420
 DB 361 PRPVQSKRFVPGCVCSERTCSNLTLTSGSKHKSFLDDDLTRLMNTEKITSCTDH 420
 QY 421 RYCQRKSYSLQVPSDILHPVELHDFSWKLVPKDRSLVLPKQLQOHTHEKPCNTSF 480
 DB 421 RYCQRKSYSLQVPSDILHPVELHDFSWKLVPKDRSLVLPKQLQOHTHEKPCNTSF 480
 QY 481 SYLVASAIPODLYFGSPCGASIKQIOVKONISVTLTFPASPFOEASRGLTVSFLPY 540
 DB 481 SYLVASAIPODLYFGSPCGASIKQIOVKONISVTLTFPASPFOEASRGLTVSFLPY 540
 QY 541 FREBGFVTFPDTSKVLTLPNMDRGPLSTVSWMISVPRDOVACLTFRERSGVVCQ 600
 DB 541 FREBGFVTFPDTSKVLTLPNMDRGPLSTVSWMISVPRDOVACLTFRERSGVVCQ 600
 QY 601 TGRAFMIIOBQRTAEIIFSLDEDLPPKSPFHHSFWNISNCSPTSGKQDLTFSVTLT 660
 DB 601 TGRAFMIIOBQRTAEIIFSLDEDLPPKSPFHHSFWNISNCSPTSGKQDLTFSVTLT 660
 QY 661 PRYVDLVIILAAVGGVLLSALGLIICVKKKKKTKNKPANCINYNN 710
 DB 661 PRYVDLVIILAAVGGVLLSALGLIICVKKKKKTKNKPANCINYNN 710
 RESULT 7
 AAM49640
 ID AAM49640 standard; protein; 749 AA.
 XX
 XX AAM49640;
 AC
 XX
 XX 17-MAY-2002 (first entry)
 DT
 XX Human tumour-associated antigen B345 protein.
 DE
 XX Tumour-associated antigen; human; B345; cytosolic; cell communication;
 KW cell interaction; signal transduction; metastasis; cancer; colon;
 KW immunotherapy; carcinoma; lung; diagnosis.
 XX
 XX Homo sapiens.
 OS

CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

CC
XX
SQ

Query Match 77.6%; Score 3395; DB 4; Length 649;

Beet Local Similarity 99.2%; Pred. No. 2.8e-288;

Matches 644; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 188 MOEGVKMALHLPFHPRVNVSFGFSIANRSSIKRLCTIESVEEGSGATLMSANYPEGFPED 247

DB 1 MOEGVKMALHLPFHPRVNVSFGFSIANRSSIKRLCTIESVEEGSGATLMSANYPEGFPED 60

QY 248 ELMTQFVVPALRLASVSFLANPNLSNCKEKREVEYITGTTNPBFLEDKQPGNNAG 307

DB 61 ELMTCQFVVPALRLASVSFLANPNLSNCKEKREVEYITGTTNPBFLEDKQPGNNAG 120

QY 308 NNFLSLGGDDADQAPGILRLQFOYLVOHPONESNKIYVVDLSNBRAMSLTEPRPVKOS 367

DB 121 NNFLSLGGDDADQAPGILRLQFOYLVOHPONESNKIYVVDLSNBRAMSLTEPRPVKOS 180

QY 368 RKFPVGCPCVLESRTCSNLTLTSGSKHISFLCDLTRLMMNVETISCTDHRVCORKS 427

DB 181 RKFPVGCPCVLESRTCSNLTLTSGSKHISFLCDLTRLMMNVETISCTDHRVCORKS 240

QY 428 YSLQVPSDILHLPVELHDPMSWGLVLPKORLSLVVPAQQLQOHTHEKPTNSFTLVASA 487

DB 241 YSLQVPGDILHLPVELHDPMSWGLVLPKORLSLVVPAQQLQOHTHEKPTNSFTLVASA 300

QY 488 IPSOOLYFGSPFCPGGSIKQIOVKONISVTLRTFAPSFQXOEAASROGLTVSFIYFKEGVF 547

DB 301 IPSOOLYFGSPFCPGGSIKQIOVKONISVTLRTFAPSFQXOEAASROGLTVSFIYFKEGVF 360

QY 548 TVTPDTKSKVYLRTPNMDRGLPSLTSSVSNISVPRDQVACLTFPKERSGVVQOTGRAPMI 607

DB 361 TVTPDTKSKVYLRTPNMDRGLPSLTSSVSNISVPRDQVACLTFPKERSGVVQOTGRAPMI 420

QY 608 IOEORTRAEITSLDEDVLPKPSFPHHSFWNVISNCSPTSGQDLILBSVTLTPTVDLT 667

DB 421 IOEORTRAEITSLDEDVLPKPSFPHHSFWNVISNCSPTSGQDLILBSVTLTPTVDLT 480

QY 668 VILIAVGGVLLLSALGLIICCVKKKKKTKNGPAVGIYNNININTEMROPKKFOKGRK 727

DB 481 VILIAVGGVLLLSALGLIICCVKKKKKTKNGPAVGIYNNININTEMROPKKFOKGRK 540

QY 728 DNDSHVAVIEDTMYGHLLODSSGFLQPEVDYTRPFGGTGVCPEPSPITCSRAPATAK 787

DB 541 DNDSHVAVIEDTMYGHLLODSSGFLQPEVDYTRPFGGTGVCPEPSPITCSRAPATAK 600

QY 788 LATEPPPPSPSESEPEPTSHPNNGDVSSKDTDILPLXQEBPMPAP 836

DB 601 LATEPPPPSPSESEPEPTSHPNNGDVSSKDTDILPLXQEBPMPAP 649

RESULT 9

ADL71669 standard; protein; 414 AA.

ADL71669;

20-MAY-2004 (first entry)

Novel human secreted protein fragment seqid 273.

antiinflammatory; neuroprotective; nootropic; antiParkinsonian;

KW anticonvulsant; antilipemic; CNS; gynaecological; antiarthritic;
KW antiasthmatic; anti-HIV; virucide; endocrine; cytostatic;
KW immunosuppressive; anti-allergic; cardiovascular; respiratory;
KW dermatological; antimicrobial; gastrointestinal; gene therapy;
KW neurodegenerative disease; behavioral disorder; inflammatory condition;
KW hyperproliferative disorder; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; metabolic disorder; Tay-Sach's disease;
KW Leash-Nyhan syndrome; reproductive disorder; immunological disorder;
KW arthritis; asthma; AIDS; endocrine disorder; immune disorder;
KW Hodgkin's lymphoma; haematopoietic disorder; muscular disorder;
KW leukaemia; autoimmune disorder; allergy; cancer; cardiovascular disorder;
KW respiratory disorder; pulmonary disorder; connective tissue disorder;
KW skin disorder; CNS disorder; congenital disorder; infectious disorder;
KW gastrointestinal disorder; human; secreted protein.

OS Homo sapiens.

PN US2004034196-A1.

PD 19-FEB-2004.

PF 27-JAN-2003; 2003US-00351334.

PR 30-JUL-1998; 98US-0094657P.

PR 05-AUG-1998; 98US-0095486P.

PR 06-AUG-1998; 98US-0095454P.

PR 06-AUG-1998; 98US-0095455P.

PR 12-AUG-1998; 98US-0096319P.

PR 29-JUL-1999; 99WO-US017130.

PR 24-JAN-2000; 2000US-00489847.

PR 25-JAN-2002; 2002US-0350898P.

XX (KOMA/) KOMATSUJIS G A.

PA (ROSE/) ROSEN C A.

PA (RUBE/) RUBEN S M.

PA (DUAN/) DUAN D R.

PA (MOOR/) MOORE P A.

PA (SHIY/) SHI Y.

PA (LAF/) LAFLEUR D W.

PA (WEIY/) WEI Y.

XX Komatsulis GA, Rosen CA, Ruben SM, Duan DR, Moore PA, Shi Y;

PI Lafleur DW, Wei Y;

XX MPI; 2004-180094/17.

DR MPI; 2004-180094/17.

XX New human secreted nucleic acid, useful for diagnosing and treating

PT neurodegenerative, inflammatory, hyperproliferative, metabolic,

PT reproductive, cardiovascular, respiratory or immunological disorders or

PT diseases.

XX Disclosure; SEQ ID NO 273; 234pp; English.

XX The invention describes an isolated human nucleic acid molecule (I)

CC comprising a polynucleotide having a nucleotide sequence at least 95%

CC identical to: a sequence polynucleotide fragment of SEQ ID NO: X or of

CC the cDNA sequence included in ATCC Deposit No: Z, which is hybridizable

CC Protein fragment of the invention. Note: This sequence does not appear in
 CC the printed specification but is available in electronic format from the
 CC US patent office at ftp.segdata.uspto.gov/segdata.html?DocID=20040034196.

XX Sequence 414 AA;

Query Match 49.2%; Score 2155; DB 8; Length 414;
 Best Local Similarity 100.0%; Pred. No. 9,4e-180;
 Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPKGAFAFIALPRESNTIVLIKLGPTTLAKPCYIV 60
 DB 1 MAGLNCGVSIALLGVLLGAARLPKGAFAFIALPRESNTIVLIKLGPTTLAKPCYIV 60
 QY 61 SKRHITMLSIKSGRIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGEVQLQPTSLPT 120
 DB 61 SKRHITMLSIKSGRIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGEVQLQPTSLPT 120
 QY 121 LNRFTIMDVKAKHSIGLELQFSIRPLRQIGPESCPDGVTHISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDVKAKHSIGLELQFSIRPLRQIGPESCPDGVTHISGRIDATVVRIGTFCSN 180
 QY 181 GTVSRIRKQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
 DB 181 GTVSRIRKQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
 QY 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKEERYEYIPGSTTNEPEVKLEDK 300
 DB 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKEERYEYIPGSTTNEPEVKLEDK 300
 QY 301 QPGNMAGNFNLISLQCCDDAOSPGILRLQFVLVQHNPONESKITYVDLSNERAMSLTTE 360
 DB 301 QPGNMAGNFNLISLQCCDDAOSPGILRLQFVLVQHNPONESKITYVDLSNERAMSLTTE 360
 QY 361 PRPVQSKRKFVPGCFVCLSESTCSNLTLTSGSKKISFLCDDLTRLMMANVEK 413
 DB 361 PRPVQSKRKFVPGCFVCLSESTCSNLTLTSGSKKISFLCDDLTRLMMANVEK 413

RESULT 10
 AAY91592
 ID AAY91592 standard; protein; 443 AA.

XX AAY91592;

DT 29-JUN-2000 (first entry)

DE Human secreted protein sequence encoded by gene 6 SEQ ID NO:265.

KM Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KM antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
 KM osteopathic; antidiabetic; antibacterial; antidiabetic; antiashma;
 KM antipruritic; cardiac; gene therapy; cancer; neurological disorder;
 KM immune disease; inflammation; blood disorder; tumour.

XX Homo sapiens.

XX WO200006698-A1.

PD 10-FEB-2000.

PF 29-JUL-1999; 99WO-US017130.

PR 30-JUL-1998; 98US-0094657P.

PR 05-AUG-1998; 98US-0095486P.

PR 06-AUG-1998; 98US-0095455P.

PR 12-AUG-1998; 98US-0095319P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Komateoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y,
 Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;

PI Soppet DR, Endress GA, Edner R, Olsen HS, Mucenaki M;
 XX WPI; 2000-195282/17.

PT New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancer, neurological
 PT disorders, immune diseases, inflammation or blood disorders.

PS Disclosure; Page 18; 634pp; English.

The polynucleotide sequences given in AAA26346 to AAA26458 encode the
 human secreted proteins given in AAY91451 to AAY91691. The human secreted
 proteins can have activities based on the tissues and cells they are
 expressed in. Examples of the activities are: cytostatic;
 immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
 antiallergic; osteopathic; antidiabetic; antibacterial; antidiabetic;
 antiashma; antipruritic; and cardiac. The polynucleotides and their
 corresponding secreted proteins are useful for preventing, treating or
 ameliorating medical conditions, e.g. by protein or gene therapy. Also
 pathological conditions can be diagnosed by determining the amount of the
 proteins in a sample or by determining the presence of mutations in the
 CC polynucleotides. Specific uses are described for each of the
 CC polynucleotides, based on which tissues they are most highly expressed
 in, and include developing products for the diagnosis or treatment of
 cancer, tumours, neurodegenerative disorders, developmental abnormalities
 and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
 CC Alzheimer's and behavioural disorders, schizophrenia, osteoporosis,
 CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
 CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
 CC reproductive disorders, gastrointestinal disorders, respiratory disorders
 CC and metabolic disorders. The proteins or polynucleotides can also be used
 CC as food additives or preservatives. The proteins are also useful for
 CC identifying their binding partners. AAA6337 to AAA26345 and AAY91450 are
 CC sequences used in the exemplification of the present invention

XX Sequence 443 AA;

Query Match 49.2%; Score 2155; DB 3; Length 443;
 Best Local Similarity 100.0%; Pred. No. 1e-179;
 Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPKGAFAFIALPRESNTIVLIKLGPTTLAKPCYIV 60
 DB 30 MAGLNCGVSIALLGVLLGAARLPKGAFAFIALPRESNTIVLIKLGPTTLAKPCYIV 89
 QY 61 SKRHITMLSIKSGRIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGEVQLQPTSLPT 120
 DB 90 SKRHITMLSIKSGRIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGEVQLQPTSLPT 149
 QY 121 LNRFTIMDVKAKHSIGLELQFSIRPLRQIGPESCPDGVTHISGRIDATVVRIGTFCSN 180
 DB 150 LNRFTIMDVKAKHSIGLELQFSIRPLRQIGPESCPDGVTHISGRIDATVVRIGTFCSN 209
 QY 181 GTVSRIRKQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
 DB 210 GTVSRIRKQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 269
 QY 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKEERYEYIPGSTTNEPEVKLEDK 300
 DB 270 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKEERYEYIPGSTTNEPEVKLEDK 329
 QY 301 QPGNMAGNFNLISLQCCDDAOSPGILRLQFVLVQHNPONESKITYVDLSNERAMSLTTE 360
 DB 330 QPGNMAGNFNLISLQCCDDAOSPGILRLQFVLVQHNPONESKITYVDLSNERAMSLTTE 389
 QY 361 PRPVQSKRKFVPGCFVCLSESTCSNLTLTSGSKKISFLCDDLTRLMMANVEK 413
 DB 390 PRPVQSKRKFVPGCFVCLSESTCSNLTLTSGSKKISFLCDDLTRLMMANVEK 442

RESULT 11
 ADL71667

ID	ADL71667	standard; protein; 443 AA.
XX		
AC	ADL71667;	
XX		
DT	20-MAY-2004	(first entry)
XX		
DE	Novel human secreted protein fragment seqid 271.	
KM	antiinflammatory; neuroprotective; nootropic; antiparkinsonian;	
KM	anticonvulsant; antihypaemic; CNS; gynaecological; antiarthritis;	
KM	antialasthmatic; anti-HIV; virucide; endocrine; cytostatic;	
KM	immunosuppressive; antiallergic; cardiovascular; respiratory;	
KM	dermatological; antimicrobial; gastrointestinal; gene therapy;	
KM	neurodegenerative disease; behavioral disorder; inflammatory condition;	
KM	hyperproliferative disorder; Alzheimer's disease; Parkinson's disease;	
KM	Huntington's disease; metabolic disorder; Tay-Sachs disease;	
KM	Leish-Nyman syndrome; reproductive disorder; immunological disorder;	
KM	arthritis; ashma; AIDS; endocrine disorder; immune disorder;	
KM	Hodgkin's lymphoma; haematopoietic disorder; muscular disorder;	
KM	leukaemia; autoimmune disorder; allergy; cancer; cardiovascular disorder;	
KM	respiratory disorder; pulmonary disorder; connective tissue disorder;	
KM	skin disorder; CNS disorder; congenital disorder; infectious disorder;	
XX	gastrointestinal disorder; human; secreted protein.	
XX		
OS	Homo sapiens.	
XX		
PN	US2004034196-A1.	
XX		
PD	19-FEB-2004.	
XX		
PF	27-JAN-2003; 2003US-00351334.	
XX		
PR	30-JUL-1998; 98US-0094657P.	
XX		
PR	05-AUG-1998; 98US-0095486P.	
XX		
PR	06-AUG-1998; 98US-0095454P.	
XX		
PR	06-AUG-1998; 98US-0095455P.	
XX		
PR	12-AUG-1998; 98US-0096519P.	
XX		
PR	29-JUL-1999; 99WO-US017130.	
XX		
PR	25-JAN-2000; 2000US-00489847.	
XX		
PR	25-JAN-2002; 2002US-0350898P.	
XX		
PA	(KOMA/) KOMATSUITS G A.	
XX		
PA	(ROSE/) ROSEN C A.	
XX		
PA	(RUSE/) RUSEN S M.	
XX		
PA	(DUAN/) DUAN D R.	
XX		
PA	(MOOR/) MOORE P A.	
XX		
PA	(SHIY/) SHI Y.	
XX		
PA	(LAFLE/) LAFLEUR D W.	
XX		
PA	(WEIY/) WEI Y.	
XX		
PI	Komatsoulis GA, Rosen CA, Ruben SM, Duan DR, Moore PA, Shi Y,	
XX	Lafleur DW, Wei Y;	
DR		
XX	WPI; 2004-180094/17.	
XX		
PT	New human secreted nucleic acid, useful for diagnosing and treating	
XX	neurodegenerative, inflammatory, hyperproliferative, metabolic,	
PT	reproductive, cardiovascular, respiratory or immunological disorders or	
XX	diseases.	
XX		
PS	Disclosure; SEQ ID NO 271; 234pp; English.	
XX		
CC	The invention describes an isolated human nucleic acid molecule (1)	
XX	comprising a polynucleotide having a nucleotide sequence at least 95%	
CC	identical to: a sequence polynucleotide fragment of SEQ ID NO: X or of	
XX	the cDNA sequence included in ATCC Deposit No: Z, which is hybridizable	
CC	to SEQ ID NO: X; or a sequence encoding a polypeptide fragment, domain or	
XX	epitope of SEQ ID NO: Y or a polypeptide sequence encoded by the cDNA	
CC	sequence included in ATCC deposit No: Z, which is hybridizable to SEQ ID	
XX	NO: X, having a biological activity. The nucleic acids and polypeptides,	
CC	pharmaceutical formulations and kits are useful in diagnosing and	
XX	treating neurodegenerative diseases states, behavioral disorders	
CC	and inflammatory conditions, hyperproliferative disorders (e.g. Alzheimer's	
XX	disease).	

	disease; parkinson's diseases or huntington's disease), metabolic disorders (e.g. Tay-Sachs's disease or leish-Nyhan syndrome), reproductive disorders (e.g. Tay-Sachs's disease or leish-Nyhan syndrome), endocrine and immune disorders (e.g. Hodgkin's lymphoma), haematopoietic or muscular disorders (e.g. leukemia), autoimmune disorders, allergy, cancer, cardiovascular, respiratory or pulmonary disorders, disorders or conditions afflicting connective tissue, skin disorders, CNS disorders, congenital disorders, infectious disorders and gastrointestinal disorders. This is the amino acid sequence of a novel human secreted protein fragment of the invention. Note: This sequence does not appear in the printed specification but is available in electronic format from the US patent office at ftp.segdata.neptco.gov/segdata.html?DocID=2040034196.
SQ	Sequence 443 AA:
Query Match	49.2%; Score 2155; DB 8; Length 443;
Best Local Similarity	100.0%; Pred. No. 1e-179;
Matches 413; Conservative	0; Mismatches 0; Indels 0; Gaps 0
DY	MAGLNGGVSIALLGYLLGAARLPARGAEAFETALPRESNITVILKGTPTLLAKPCYIV 60
DB	MAGLNGGVSIALLGYLLGAARLPARGAEAFETALPRESNITVILKGTPTLLAKPCYIV 89
OY	SKRHITMLSSIGSERIIVTFPSCQSEPNHFVIEIOKNIDCMGSGCPFGVEVOLPSTSLPT 120
DB	SKRHITMLSSIGSERIIVTFPSCQSEPNHFVIEIOKNIDCMGSGCPFGVEVOLPSTSLPT 149
OY	LNRRTIMVVKAHKSIGLELOFSIPRLROI GPESCCPDGYTHSISGRIDATVARIGFCSN 180
DB	LNRRTIMVVKAHKSIGLELOFSIPRLROI GPESCCPDGYTHSISGRIDATVARIGFCSN 209
OY	GTVSRIRKMGEGVMALHLPMFPFRNVSGPSINRSIKELCTIESVFEEGGSAITMSANY 240
DB	GTVSRIRKMGEGVMALHLPMFPFRNVSGPSINRSIKELCTIESVFEEGGSAITMSANY 269
OY	PEGFEDELMTMQFVPVPAHRAVSFLFNLTNSCEKKEKVRYTTPGSTTNPEVFTLEDK 300
DB	PEGFEDELMTMQFVPVPAHRAVSFLFNLTNSCEKKEKVRYTTPGSTTNPEVFTLEDK 329
OY	QPGNMAGNFNLTLQGCDODDAPGILRILOFYLVPHONESNKIIYVVDSLNERAMSLTTE 360
DB	QPGNMAGNFNLTLQGCDODDAPGILRILOFYLVPHONESNKIIYVVDSLNERAMSLTTE 389
OY	PBPVQSKRFKPVDFGCVCLESFRCSSNLTLTGSSKHKISTFLCDDLTRLMMANVER 413
DB	PBPVQSKRFKPVDFGCVCLESFRCSSNLTLTGSSKHKISTFLCDDLTRLMMANVER 442
RESULT 12	
AU012252	ID AU012252 standard; protein; 343 AA.
XX AC	AAU012252;
XX DT	24-OCT-2001 (first entry)
DE XX	Human PRO5773 polypeptide sequence.
KW KW	Human secretary and transmembrane; PRO; mammalian; cancer; lung; breast; prostate; cervical; tumor necrosis factor-alpha; TNF-alpha; cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte; A-peptide; factor VIIA; gene therapy.
OS OS	Homo sapiens.
PN PN	WO200140466-A2.
XX PD	07-JUN-2001.
PF PF	01-DEC-2000; 2000MO-USO32678.
XX PR	01-DEC-1999; 99WO-USO28301.
PR PR	01-DEC-1999; 99WO-USO28634.

PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028554.
 PR 09-DEC-1999; 99WO-US028565.
 PR 09-DEC-1999; 99US-0170262P.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 20-DEC-1999; 99WO-US030999.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US000365.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 03-MAR-2000; 2000US-0187202P.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 PR 21-MAR-2000; 2000WO-US007532.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 12-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 05-JUN-2000; 2000US-0209832P.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 11-AUG-2000; 2000WO-US022031.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 10-NOV-2000; 2000WO-US030873.
 XX
 PA (GERTH) GENENTECH INC.
 XX
 PI Baker KP, Barresi M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerlesen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI; 2001-408281/43.
 DR N-PSDB; AAS21324.
 XX
 PT Isolated , secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing PRO
 PT polypeptides, and detect the presence of mammalian tumors e.g. lung,
 PT breast, prostate, cervical.
 XX
 PS Claim 12; Fig 162; 813pp; English.
 XX
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane PRO
 CC polypeptides. The PRO polypeptides are useful to detect other PRO
 CC polypeptides, to link bioactive molecules to cells expressing PRO
 CC polypeptides, to modulate biological activities of cells expressing PRO
 CC polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample. Some
 CC of the 275 sequences are also useful to stimulate the release of tumour
 CC neurosis factor-alpha (TNF-alpha) from human blood, the proliferation or
 CC differentiation of chondrocytes, the proliferation or gene expression in
 CC pericyte cells, the release of proteoglycans from cartilage, the
 CC proliferation of inner ear utricular supporting cells or of T-
 CC lymphocytes, the release of a cytokine from peripheral blood monocytes
 CC (pBMCs), or the proliferation of endothelial cells. Some of the PRO
 CC polypeptides may modulate glucose or free fatty acid uptake by skeletal
 CC muscle cells or by adipocytes; or inhibit binding of A-peptide to factor
 CC VIIA. The PRO polypeptides can be used in assays to identify molecules
 CC involved in binding interactions. The polynucleotides encoding PRO
 CC polypeptides can be used to generate probes, antisense RNA/DNA,

CC transgenic or knock out animals and can be used in gene therapy
 XX
 SQ Sequence 343 AA;
 Query Match 40.6%; Score 1779; DB 4; Length 343;
 Best Local Similarity 99.4%; Pred. No. 7e-147;
 Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGLNGVSIALLGVLGGAARLRGADEAFIALPRESNTIVLKLGTPTLLAKPCYIVI 60
 DB 1 MAGLNGVSIALLGVLGGAARLRGADEAFIALPRESNTIVLKLGTPTLLAKPCYIVI 60
 QY 61 SKRIITMLSTKSGRIYFTSCQSPENHFVIEIKNIDCMGSPCPFGVQLQSTSLIPT 120
 DB 61 SKRIITMLSTKSGRIYFTSCQSPENHFVIEIKNIDCMGSPCPFGVQLQSTSLIPT 120
 QY 121 LNRFTFMDVKAKHSIGLELOPSIRLRQIGGESCPCGVTHSIGRIDATVVRIGTFCSN 180
 DB 121 LNRFTFMDVKAKHSIGLELOPSIRLRQIGGESCPCGVTHSIGRIDATVVRIGTFCSN 180
 QY 181 GTVSRIKMGEGVKMALHPMFHPNVSGFSIANRSLIKRLCIISVFEGBGSAITMSANY 240
 DB 181 GTVSRIKMGEGVKMALHPMFHPNVSGFSIANRSLIKRLCIISVFEGBGSAITMSANY 240
 QY 241 PEGFPEDELMTWQVVAHLRASVSLNPNLSNCRKEERYEYIIPGTTPEYFKLEDK 300
 DB 241 PEGFPEDELMTWQVVAHLRASVSLNPNLSNCRKEERYEYIIPGTTPEYFKLEDK 300
 QY 301 QPGMAGNFNLSLQGCDDQASPGIILQFOVLVQHONESNK 343
 DB 301 QPGMAGNFNLSLQGCDDQASPGIILQFOVLVQHONESSE 343
 RESULT 13
 ABO17696
 ID ABO17696 standard; protein; 343 AA.
 XX
 AC ABO17696;
 XX
 DT 26-AUG-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO5773.
 XX
 KW Human; secreted and transmembrane protein; PRO; antiinflammatory;
 KW antiarteriosclerotic; cardiatic; anti-infertility; anti-HIV; cytosatic;
 KW antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;
 KW TNF-alpha release; cell proliferation; cell differentiation;
 KW gene expression modulator; proteoglycan release; cytokine release;
 KW tumour; inflammatory disease; organ failure; atherosclerosis;
 KW cardiac injury; infertility; birth defect; premature aging; AIDS;
 KW acquired immunodeficiency syndrome; cancer; diabetic complication;
 KW chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;
 KW bioreactor; tissue typing.
 XX
 OS Homo sapiens.
 XX
 PN US2003032156-A1.
 XX
 PD 13-FEB-2003.
 XX
 PF 06-MAY-2002; 2002US-00140474.
 XX
 XX 31-MAR-1997; 97WO-US005230.
 PR 12-JUN-1998; 98WO-US012456.
 PR 14-JUL-1998; 98WO-US014552.
 PR 28-AUG-1998; 98WO-US017888.
 PR 10-SEP-1998; 98WO-US018824.
 PR 14-SEP-1998; 98WO-US019093.
 PR 14-SEP-1998; 98WO-US019094.
 PR 16-SEP-1998; 98WO-US019177.
 PR 16-SEP-1998; 98WO-US019330.
 PR 17-SEP-1998; 98WO-US019437.
 PR 07-OCT-1998; 98WO-US021141.

```

PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005130.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010723.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021030.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028554.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747253.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.

```

```

PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
PR (GETH ) GENENTECH INC.
PR PA
PR Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PR Gerltsen ME, Goddard A, Godowski PJ, Gurney AB, Sherwood S;
PR Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
PR WPI; 2003-341980/32.
PR DR N-PSDB; ACD23933.
PR DR
PR PT New secreted and transmembrane PRO nucleic acids, for treating
PR PT inflammation, organ failure, atherosclerosis, cardiac injury,
PR PT infertility, birth defects, premature aging, acquired immunodeficiency
PR PT syndrome (AIDS), or cancer.
PR XX
PR PS Claim 12; Fig 162; 660pp; English.
PR XX
PR CC The invention describes an isolated nucleic acid (I) comprising, or which
PR CC has 80 % sequence identity to, or the full-length coding sequence of, one
PR CC of 275 nucleotide sequences, and which encodes a corresponding
PR CC polypeptide selected from 275 amino acid sequences, where all sequences
PR CC are given in the specification. The polypeptide encoded by (I) is used to
PR CC detect PRO polypeptides, link a bioactive molecule to a cell expressing a
PR CC PRO polypeptide, modulate a biological activity of a cell, stimulate the
PR CC release of tumour necrosis factor (TNF)-alpha from human blood, modulate
PR CC the uptake of glucose or free fatty acid by cells, stimulate or inhibit
PR CC the proliferation or differentiation of cells or gene expression,
PR CC stimulate the release of proteoglycans, stimulate the release of cytokine
PR CC from peripheral blood mononuclear cells, inhibit the binding of A-peptide
PR CC to factor VIIa, or detect the presence of tumour in a mammal. The nucleic
PR CC acid and polypeptide encoded by it, are useful for treating inflammatory
PR CC diseases, organ failure, atherosclerosis, cardiac injury, infertility,
PR CC birth defects, premature aging, acquired immunodeficiency syndrome
PR CC (AIDS), cancer, or diabetic complications. The nucleic acid is useful as
PR CC hybridisation probes, in chromosome and gene mapping, and in generating
PR CC antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,
PR CC diagnostics, biosensors or bioreactors. Both are useful in tissue typing.
PR CC This is the amino acid sequence of a novel human secreted and
PR CC transmembrane PRO polypeptide
PR XX
PR SQ Sequence 343 AA;

```

```

Query Match 40 6%; Score 1779; DB 6; Length 343;
Best Local Similarity 99.4%; Pred. No. 7e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAGLNCGVSIALLGTLGARLPRGAEAFIALPRESNITVLKLGPTLLAKRCYVI 60
DB 1 MAGLNCGVSIALLGTLGARLPRGAEAFIALPRESNITVLKLGPTLLAKRCYVI 60
QY 61 SKRHITMLSIKSGERIVTFSCQSPENHFVIEIQKNDICMGSCPCPFGEVQLOPSTSLPT 120
DB 61 SKRHITMLSIKSGERIVTFSCQSPENHFVIEIQKNDICMGSCPCPFGEVQLOPSTSLPT 120
QY 121 LNRFTIMDKAKHSIGLEQFSIPRLRIGSGESCPDQVTHSISGRIDATVVRIGTFCPSN 180
DB 121 LNRFTIMDKAKHSIGLEQFSIPRLRIGSGESCPDQVTHSISGRIDATVVRIGTFCPSN 180

```


QY 181 GTVSRIRKQEGVKMLHLPMFHPRVSGFSIANSRIRLCITIESEFEGEGSATLMSANY 240
Db 181 GTVSRIRKQEGVKMLHLPMFHPRVSGFSIANSRIRLCITIESEFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTQMFVPAHLRASVSFLNFINLSNCRKEERYEYYPGSTTNEVFKEBDK 300
Db 241 PEGFPEDELMTQMFVPAHLRASVSFLNFINLSNCRKEERYEYYPGSTTNEVFKEBDK 300
QY 301 QPGMAGNFNLSLQCCDDAOSPGLRLQFQVLYVQHPONESNK 343
Db 301 QPGMAGNFNLSLQCCDDAOSPGLRLQFQVLYVQHPONESNK 343
RESULT 14
ID ABU80950 standard; protein, 343 AA.
XX
AC ABU80950;
XX
DT 23-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #81.
XX
KM Human; PRO polypeptide; secreted and transmembrane protein;
KM anti-PRO antibody; diagnostic assay; gene expression; diabetes;
KM bone disorder; cartilage disorder; rheumatoid arthritis; obesity;
KM sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;
KM hearing loss; coagulation disorder; stroke; heart attack; cardiacnt;
KM antidiabetic; anorectic; vulnerability; antiarthritis; osteopathic;
KM antirheumatic; auditory; cerebroprotective; angiogenic.
XX
OS Homo sapiens.
XX
PN US2003004311-A1.
XX
PD 02-JAN-2003.
XX
PF 19-DEC-2001; 2001US-00028072.
XX
PR 18-JUN-1997; 97US-0049911P.
PR 26-AUG-1997; 97US-0056974P.
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 19-SEP-1997; 97US-0059352P.
PR 19-SEP-1997; 97US-0059588P.
PR 24-SEP-1997; 97US-0059836P.
PR 17-OCT-1997; 97US-0062250P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 17-OCT-1997; 97US-0063756P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063082P.
PR 27-OCT-1997; 97US-0063127P.
PR 27-OCT-1997; 97US-0063322P.
PR 28-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063561P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063733P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 12-NOV-1997; 97US-0065846P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.

PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 11-DEC-1997; 97US-0069212P.
PR 11-DEC-1997; 97US-0069278P.
PR 11-DEC-1997; 97US-0069334P.
PR 16-DEC-1997; 97US-0069694P.
PR 23-JAN-1998; 98US-0072320P.
PR 04-FEB-1998; 98US-0073612P.
PR 09-FEB-1998; 98US-0074086P.
PR 12-MAR-1998; 98US-0077791P.
PR 20-MAR-1998; 98US-0078910P.
PR 25-MAR-1998; 98US-0079294P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079728P.
PR 31-MAR-1998; 98US-0080165P.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 16-SEP-1998; 98WO-US019177.
PR 17-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98WO-US019437.
PR 29-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 08-SEP-1999; 99WO-US020111.
PR 13-SEP-1999; 99WO-US020594.
PR 15-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 05-OCT-1999; 99WO-US021547.
PR 29-NOV-1999; 99WO-US023089.
PR 30-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028651.
PR 02-DEC-1999; 99WO-US028654.
PR 16-DEC-1999; 99WO-US028655.
PR 20-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 30-DEC-1999; 99WO-US030999.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 11-FEB-2000; 2000WO-US000376.
PR 18-FEB-2000; 2000WO-US000355.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
XX
XX (GENTH) GENENTECH INC.
XX
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
PI Geritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

DR WPI: 2003-352836/33.
DR N-PSDB; A667074.
PT New isolated PRO polypeptide useful for treating diabetes, rheumatoid
PT arthritis, sports injuries, obesity, hearing loss in mammals, stroke, or
PT heart attack.
XX
PS Claim 12; Fig 162; 643pp; English.
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The PRO
CC polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides and polynucleotides are useful for preparing a medicament
CC useful in the treatment of diabetes, bone and/or cartilage disorders
CC (e.g. rheumatoid arthritis, sports injuries, osteoarthritis), obesity,
CC hyper- or hypo-insulinaemia, hearing loss, and coagulation disorders
CC (e.g. stroke, heart attack). Anti-PRO antibodies are useful in diagnostic
CC assays for PRO, by detecting its expression in specific cells, tissues or
CC serum, and for affinity purification of PRO from recombinant cell culture
CC or natural sources. AB060870-AB081144 represent the human PRO
CC polypeptides of the invention. Note: The sequence data for this patent
CC was obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/psipdbidentity.html
XX
SQ Sequence 343 AA;

Query Match 40.6%; Score 1779; DB 6; Length 343;
Best Local Similarity 99.4%; Pred. No. 76-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGSIALGLVLLGAARLPFGAFAFEIALPRESNITVLIKIGTPTLLAKPCYIV 60
DB 1 MAGLNCGSIALGLVLLGAARLPFGAFAFEIALPRESNITVLIKIGTPTLLAKCYIV 60
QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQNIDCMGCPFGVEVOLQPTSLAPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQNIDCMGCPFGVEVOLQPTSLAPT 120
QY 121 LNRFTIWMVKAKHSIGLELOPSIFRLROI GPESCPDGVTHISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWMVKAKHSIGLELOPSIFRLROI GPESCPDGVTHISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESEFEGEGSATLMSANY 240
DB 181 GTVSRIKMQEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESEFEGEGSATLMSANY 240
QY 241 PEGFPEDEILMTQFVVPALRASVSFLNFNLSNCRKEERVEYYIPGSTTNPVEFKLEDK 300
DB 241 PEGFPEDEILMTQFVVPALRASVSFLNFNLSNCRKEERVEYYIPGSTTNPVEFKLEDK 300
QY 301 QPQNMAGNPNLSLQCCDDAOSPGILRLQFOVLVQHPPONESMK 343
DB 301 QPQNMAGNPNLSLQCCDDAOSPGILRLQFOVLVQHPPONESSE 343

RESULT 15
AB066650
ID AB066650 standard; protein; 343 AA.
XX
AC AB066650;
XX
DT 23-MAY-2003 (first entry)
XX
DE Human PRO polypeptide #81.
XX
KW Human; PRO polypeptide; secreted and transmembrane protein;
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;
XX differential; chondrocyte; tumour; genetic disorder; cytostatic.
OS Homo sapiens.
XX
PN US2003036180-A1.
XX

PD 20-FEB-2003.
XX
PF 09-MAY-2002; 2002US-00143114.
XX
PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019053.
PR 14-SEP-1998; 98WO-US019094.
PR 16-SEP-1998; 98WO-US019177.
PR 17-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98WO-US021147.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005150.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020584.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028554.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030911.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.

PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00786498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001US-00806666.
PR 09-MAR-2001; 2001WO-US0082706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 05-JUN-2001; 2001WO-US017800.
PR 14-JUN-2001; 2001US-00874503.
PR 19-JUN-2001; 2001US-00882636.
PR 20-JUN-2001; 2001US-00883432.
PR 21-JUN-2001; 2001WO-US019692.
PR 22-JUN-2001; 2001US-00887879.
PR 29-JUN-2001; 2001WO-US020116.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerltsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2003-332040/31.
DR N-PSDB; ACA03683.
XX
PT New secreted and transmembrane PRO nucleic acids, useful for gene
PT therapy, in chromosome and gene mapping, as chromosome markers, in tissue
PT typing, and in chromosome identification.
XX
PS Claim 12; Fig 162; 660pp; English.
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The PRO
CC polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides are useful for detecting other PRO polypeptides, for linking
CC bioactive molecules to cells expressing PRO polypeptides, for modulating
CC biological activities of cells expressing PRO polypeptides, and for
CC identifying agonists or antagonists. The PRO polypeptides are useful for
CC for stimulating the release of tumour necrosis factor (TNF)-alpha from
CC human blood, for stimulating the proliferation or differentiation of
CC chondrocytes, and detecting the presence of tumours. The polynucleotide
CC sequences encoding PRO polypeptides are useful as hybridisation probes,
CC in chromosome and gene mapping, in the generation of antisense RNA and
CC DNA, in the preparation of PRO polypeptides, for generating transgenic
CC animals or knockout animals, for the genetic analysis of individuals with
CC genetic disorders, and in gene therapy. ABU6570-ABU6844 represent the
CC human PRO polypeptides of the invention. Note: The sequence data for this
CC patent was obtained in electronic format directly from the USPRO web site
CC at seqdata.uspto.gov/patidbIdentify.html
XX
SQ Sequence 343 AA;

Query Match 40.6%; Score 1779; DB 6; Length 343;
Best Local Similarity 99.4%; Pred. No. 7e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCVSTALLGVLLGGAARLRPGAAFEIALPRESNITVLKLGPTLLAKPCYIV 60
|||||

Db 1 MAGLNCVSTALLGVLLGGAARLRPGAAFEIALPRESNITVLKLGPTLLAKPCYIV 60
QY 61 SKRHITMLSTIKSGRIVFTFSQSPENHFVIEIQKNIDCMGSPCFGEVQIQPSTSLPT 120
Db 61 SKRHITMLSTIKSGRIVFTFSQSPENHFVIEIQKNIDCMGSPCFGEVQIQPSTSLPT 120
QY 121 LNRFTIWDVKAHKSIGLEQFSIPRLROIIGGESCPCGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTIWDVKAHKSIGLEQFSIPRLROIIGGESCPCGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIMQEGVKAHLALPMFHPNVSQFSIANRSSIKRLCTIESVEEGSATILMSANY 240
Db 181 GTVSRIMQEGVKAHLALPMFHPNVSQFSIANRSSIKRLCTIESVEEGSATILMSANY 240
QY 241 PEGFPEDELMTWQFVPAHLRASVSLFNFLNSCERKEERYEYIIPGSTNPEVFKLEDK 300
Db 241 PEGFPEDELMTWQFVPAHLRASVSLFNFLNSCERKEERYEYIIPGSTNPEVFKLEDK 300
QY 301 QPQMAGNFNLSLOGCDQDQSPGILRLQFVLYQHQNESNK 343
Db 301 QPQMAGNFNLSLOGCDQDQSPGILRLQFVLYQHQNESSE 343

Search completed: May 4, 2006, 22:54:39
Job time : 138 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 22:25:30 ; Search time 47 Seconds
(without alignments)

1470.572 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4377
Sequence: 1 MAGLNCGVSTALLGVLLGA.....SSKDTDIPLXTQRPMPBAE 836

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/6.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/H.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/RE.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3625.5	82.8	709	2	US-09-489-847-132 Sequence 132, App
2	2155	49.2	414	2	US-09-489-847-273 Sequence 273, App
3	2155	49.2	443	2	US-09-489-847-271 Sequence 271, App
4	827	18.9	159	2	US-09-489-847-272 Sequence 272, App
5	141	3.2	3623	2	US-09-341-461-2 Sequence 2, Appli
6	126.5	2.9	590	2	US-09-902-540-11386 Sequence 11386, A
7	118.5	2.7	2476	1	US-08-276-967-2 Sequence 2, Appli
8	118	2.7	2766	2	US-09-964-956-62 Sequence 62, Appli
9	114	2.6	1589	2	US-08-755-587-189 Sequence 189, App
10	112.5	2.6	440	2	US-09-538-092-599 Sequence 999, App
11	112.5	2.6	451	2	US-09-949-016-9282 Sequence 9282, Ap
12	112.5	2.6	1694	2	US-09-560-385A-12 Sequence 12, Appli
13	112.5	2.6	1725	2	US-09-560-385A-10 Sequence 10, Appli
14	108	2.5	699	2	US-09-949-016-6138 Sequence 6138, Ap
15	107.5	2.5	649	2	US-09-248-796A-20641 Sequence 20641, A
16	107.5	2.5	744	2	US-09-854-845-43 Sequence 43, Appli
17	107.5	2.5	771	2	US-09-854-845-47 Sequence 47, Appli
18	107.5	2.5	843	2	US-09-854-845-27 Sequence 27, Appli
19	107.5	2.5	870	2	US-09-854-845-31 Sequence 31, Appli
20	107.5	2.5	1719	1	US-08-459-568-4 Sequence 4, Appli
21	107.5	2.5	1719	1	US-08-399-411-4 Sequence 4, Appli
22	107.5	2.5	1719	2	US-08-516-859A-4 Sequence 4, Appli
23	107.5	2.5	1719	2	US-09-586-472-4 Sequence 4, Appli
24	107.5	2.5	1719	2	US-09-528-706-4 Sequence 4, Appli
25	107.5	2.5	1719	2	US-10-024-450-4 Sequence 4, Appli
26	107	2.4	739	2	US-09-854-845-45 Sequence 45, Appli
27	107	2.4	766	2	US-09-854-845-49 Sequence 49, Appli

ALIGNMENTS

```
RESULT 1
; Sequence 132, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT FILING DATE: US/09/489, 847
; EARLIER FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 132
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (189)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (275)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (414)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (438)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (641)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
```

```
LOCATION: (643)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (696)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (697)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-489-847-132

Query Match      82.8%; Score 3625.5; DB 2; Length 709;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 698; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 1 MAGINCGVSIALGLVLLGAAARLPRGAAFEIALPRESNITVLIKLGPTLLAKPCYIV 60
DB 1 MAGINCGVSIALGLVLLGAAARLPRGAAFEIALPRESNITVLIKLGPTLLAKPCYIV 60
QY 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQNIDCMGSPCPFGVQLQPTSLPT 120
DB 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQNIDCMGSPCPFGVQLQPTSLPT 120
QY 121 LNRTFIMDVKAHKSIGLELQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIMDVKAHKSIGLELQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIRKLCIISVFEGESATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIRKLCIISVFEGESATLMSANY 240
QY 241 PEGPEDELMTWQFVPAHLRASAVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEVKLEDK 300
DB 241 PEGPEDELMTWQFVPAHLRASAVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEVKLEDK 300
QY 241 PEGPEDELMTWQFVPAHLRASAVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEVKLEDK 300
DB 241 PEGPEDELMTWQFVPAHLRASAVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEVKLEDK 300
QY 301 QPGMAGNFNLSLOGCDQDQSPGILRLQFOVLYQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGNFNLSLOGCDQDQSPGILRLQFOVLYQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKQSRKFPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDLTRLMMNVEKTSCTDH 420
DB 361 PRPVKQSRKFPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDLTRLMMNVEKTSCTDH 420
QY 421 RYCCOKKSYSLQVPSDILHLPELHDFSMKLVLPKORLSLVLPACKLQOHTHEKFCNTSF 480
DB 421 RYCCOKKSYSLQVPSDILHLPELHDFSMKLVLPKORLSLVLPACKLQOHTHEKFCNTSF 480
QY 481 SYLVASAIPOSQDLYFGSFCPGGSIKQIQVKONISVTLRTFAPSFXQEASROGLTVSFIPY 540
DB 481 SYLVASAIPOSQDLYFGSFCPGGSIKQIQVKONISVTLRTFAPSFXQEASROGLTVSFIPY 540
QY 541 FKKEGVFVTVPTDYSKYVLRTPNMDRGIPSLTSVSWNISVPRDQVACLTFFKERSGVVQ 600
DB 541 FKKEGVFVTVPTDYSKYVLRTPNMDRGIPSLTSVSWNISVPRDQVACLTFFKERSGVVQ 600
QY 601 TGRAFMIIQEOBRTAREEIFSLEDEVLPRKSPFHHSFWNISCSCTSGKOLDLFSVLT 660
DB 601 TGRAFMIIQEOBRTAREEIFSLEDEVLPRKSPFHHSFWNISCSCTSGKOLDLFSVLT 660
QY 661 PRYVDLTVILIAAVGGVLLSALGLIICVAKKKKTKNKGPVAVGINXN 710
DB 661 PRYVDLTVILIAAVGGVLLSALGLIICVAKKKKTKNKGPVAVGINXN 710

RESULT 2
US-09-489-847-273
; Sequence 273, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
```

```
CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 273
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-273

Query Match      49.2%; Score 2155; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 5,4e-205;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGINCGVSIALGLVLLGAAARLPRGAAFEIALPRESNITVLIKLGPTLLAKPCYIV 60
DB 1 MAGINCGVSIALGLVLLGAAARLPRGAAFEIALPRESNITVLIKLGPTLLAKPCYIV 60
QY 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQNIDCMGSPCPFGVQLQPTSLPT 120
DB 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQNIDCMGSPCPFGVQLQPTSLPT 120
QY 121 LNRTFIMDVKAHKSIGLELQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIMDVKAHKSIGLELQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIRKLCIISVFEGESATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIRKLCIISVFEGESATLMSANY 240
QY 241 PEGPEDELMTWQFVPAHLRASAVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEVKLEDK 300
DB 241 PEGPEDELMTWQFVPAHLRASAVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEVKLEDK 300
QY 241 PEGPEDELMTWQFVPAHLRASAVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEVKLEDK 300
DB 241 PEGPEDELMTWQFVPAHLRASAVSFLNFNLSNCRKEEREVEYYIPGSTTNPVEVKLEDK 300
QY 301 QPGMAGNFNLSLOGCDQDQSPGILRLQFOVLYQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGNFNLSLOGCDQDQSPGILRLQFOVLYQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKQSRKFPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDLTRLMMNVEK 413
DB 361 PRPVKQSRKFPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDLTRLMMNVEK 413

RESULT 3
US-09-489-847-271
; Sequence 271, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
```

```

; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 271
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-271

Query Match      49.2%; Score 2155; DB 2; Length 443;
Best Local Similarity 100.0%; Pred. No. 6,1e-205;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAGNCGVSIALGLVLLGAARLPKGAFAFEIALPRESNITVLIKLGPTTLAKPCYVI 60
        |||
DB      30 MAGNCGVSIALGLVLLGAARLPKGAFAFEIALPRESNITVLIKLGPTTLAKPCYVI 89

QY      61 SKRIITMISTKSGRIYVTFSCQSPENHFVIEIQKIDCMGCPGFEVQLQPTSLPT 120
        |||
DB      90 SKRIITMISTKSGRIYVTFSCQSPENHFVIEIQKIDCMGCPGFEVQLQPTSLPT 149

QY      121 LNRFTMDVKAKHSIGLEQSIPLRLQIGGSCPDGVTSHISGRIDATVVRIGTFCSN 180
        |||
DB      150 LNRFTMDVKAKHSIGLEQSIPLRLQIGGSCPDGVTSHISGRIDATVVRIGTFCSN 209

QY      181 GTVSRIKMQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIEVFEGSGSATLMSANY 240
        |||
DB      210 GTVSRIKMQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIEVFEGSGSATLMSANY 269

QY      241 PEGPEDELMTWQVVPVPAHLRASVSFLNPNLSNCRKEERYEYIYPGSTNPEVFKEDK 300
        |||
DB      270 PEGPEDELMTWQVVPVPAHLRASVSFLNPNLSNCRKEERYEYIYPGSTNPEVFKEDK 329

QY      301 QPGMAGNFNLISLQCDQDQSPGILRLQFVLVQHPNENKITYVVDLSERAMSLTIE 360
        |||
DB      330 QPGMAGNFNLISLQCDQDQSPGILRLQFVLVQHPNENKITYVVDLSERAMSLTIE 389

QY      361 PRPVKQSRKFPVPGCFVLESRTCSNLTITSGSKHISFLCDDLTRLMNVK 413
        |||
DB      390 PRPVKQSRKFPVPGCFVLESRTCSNLTITSGSKHISFLCDDLTRLMNVK 442

RESULT 4
US-09-489-847-272
; Sequence 272, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 272
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

US-09-489-847-272

Query Match      18.9%; Score 827; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 1e-73;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30 FEIALPRESNITVLIKLGPTTLAKPCYIVSKRIITMISTKSGRIYVTFSCQSPENHF 89
        |||
DB      1 FEIALPRESNITVLIKLGPTTLAKPCYIVSKRIITMISTKSGRIYVTFSCQSPENHF 60

QY      90 VIEIQKIDCMGCPGFEVQLQPTSLPTLNRFTMDVKAKHSIGLEQSIPLRLQI 149
        |||
DB      61 VIEIQKIDCMGCPGFEVQLQPTSLPTLNRFTMDVKAKHSIGLEQSIPLRLQI 120

QY      150 GPGSCPDGVTSHISGRIDATVVRIGTFCSNGTVSRIM 188
        |||
DB      121 GPGSCPDGVTSHISGRIDATVVRIGTFCSNGTVSRIM 159

RESULT 5
US-09-341-461-2
; Sequence 2, Application US/09341461
; Patent No. 6586389
; GENERAL INFORMATION:
; APPLICANT: Hammond, Timothy G.
; APPLICANT: Verroust, Pierre J.
; TITLE OF INVENTION: Cubilin Protein, DNA Sequences Encoding Cubilin
; FILE REFERENCE: D6148
; CURRENT APPLICATION NUMBER: US/09/341,461
; CURRENT FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: PCT/US99/01259
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 40
; SEQ ID NO 2
; LENGTH: 3623
; TYPE: PRT
; ORGANISM: rat
; FEATURE:
; OTHER INFORMATION: amino acid sequence of rat cubilin protein
US-09-341-461-2

Query Match      3.2%; Score 141; DB 2; Length 3623;
Best Local Similarity 18.9%; Pred. No. 0.0035;
Matches 164; Conservative 85; Mismatches 315; Indels 306; Gaps 36;

QY      70 IKSGRIYVTFSCQSPENHFVIEIQKIDCMGCPGFEVQLQPTSLPTLNRFTI 126
        |||
DB      2306 VSSRRRIYLRKFTGTGSSYMKFKAQYIASCGTIVS-GDSGVISIGVPTLPYANNVCQ 2364

QY      127 WDVAHKSIGLEQSIPLRLQIGGSCPDGVT---HSISGRIDATVVRIGTFCSNGT 182
        |||
DB      2365 WPIRLPGGYLTLSFEDFNLT-QSSPG--CTKDQFVIEWNGTSGHVR-----LGRYCGNST 2415

QY      183 VSRIKMQEGVKALHLPWFHPRNVSGFSIANRSSIKRLCTIEVFEGSGSATLMSANYPE 242
        |||
DB      2416 PSSVDTSSNV-ASVKFTVDGSVTASGRLOPKSS-RQVCGDDL--GGPTGTFSPNVPN 2470

QY      243 GPFDELMTWQVVPVPAHLRASVSFLNPNLSN-----CERK 277
        |||
DB      2471 PNPGARICENTITVQEGRRIVLTFTNLRLSTQPSCSNBSGLIVPNGISNSPFLQKLSRV 2530

QY      278 EERVEYIYPGSTNPEVFKEDKQ-----GNMAGNFNLISLQCDQDQSPGI 325
        |||
DB      2531 NVTNPFKSSGNTMK-VVFTFDGSRPYGFTASYSTEDAVCGFLPSVSG--GNFSSPGY 2587

QY      326 -----LRQFVLVQHPNENKITYVVDLSERAMSLTIEPRPVKQSRKFPVPGCFV 376
        |||
DB      2588 NGIRDVARNIDCEWTLSPNRENSISITFLELISEGQDCT-----FD 2631

QY      377 CLESRTCSNLTITSGSKHISFLCDDLTRLMNVKTICTDHRVCORKSYSIQVPSDI 436
        |||
DB      2632 VLEFRVVDADQGPL-----IEK-----FC-----SLSATPAP 2657
```


Db 172 RQGNKNTFI---QPAGPFGISLNGBGHYIFLETDFKFSQAGSFRVSRPFCAP--AVICV 226
Qy 589 TFFERSGVVCGTGRAPMIIOEORTRAEIR-----SLDEDVLPK-PSFHHHSRWV 638
Db 227 TFTYHMYGL-----GQGTKLRLGSPAGSPSSLMERVGQSE-----WL 268
Qy 639 NISNCSPTSGKQ-LDLLFSVTLPRTVDLTVILIAAVGGVLLLSALGLIICVKKKKK 697
Db 269 NTSVTISGHQHPQOLIFE-----AVKGTNPAFVALGFLV-----304
Qy 698 TNKCPAVGIYXNINTEMPPROPKFKQGRKNDSHVAV-IEDTMVYGHLLQDSSGSEFLQ 756
Db 305 INHGTGSESETSVSTEKVPAPTE---KPTVPSEIYITPTEKPMVH-----MEK 350
Qy 757 PEVDTPPFGCTMGVCPSPPTICSRAPTAK-----LATEBP-----PRSPSESES 803
Db 351 PIYHTEKP---TVPTKEKPTIPTEKSTVPTKPTVKEPTLPPEGPVPAERPTTPPEGA 407
Qy 804 EP-----YTFSHPNNGDVSSKDDIP-----LLXTQEPMEPAE 836
Db 408 VPPKGPVLTVMPTSHTEKSTVHTEKPLPTGKSTIPEKPMVPTK 453

RESULT 8
US-09-964-956-62
; Sequence 62, Application US/09964956
; Patent No. 6875570
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Miller, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grose, William M
; APPLICANT: Alsebrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Bugess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shinkels, Richard A
; TITLE OR INVENTION: No. 6875570el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/09/964, 956
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235, 631
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235, 633
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235, 808
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236, 064
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236, 065
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236, 066
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236, 135
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237, 434
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238, 321
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238, 399
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238, 396
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/276, 667
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/294, 823

; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/304, 868
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 2766
; TYPE: PR
; ORGANISM: Rattus norvegicus
US-09-964-956-62

Query Match 2.7%; Score 118; DB 2; Length 2766;
Best Local Similarity 19.0%; Pred. No. 0.42; Mismatches 265; Indels 238; Gaps 34;
Matches 140; Conservative 95;

Qy 144 PRLEQIGPESCPDGVTHSISGRIDATVIRIGTCNGTVSRIRIMQGVKALHPWHP 203
Db 1865 KPPQCVPEKPPQALGSL--RTSADTSIRFTSPITSPKLPQGANRPFH-----1917
Qy 204 RNVGFSIANRSSIKRLCITIEVPEGSGATLMSANYPEGPEDEL--MTQFVVPALR 261
Db 1918 -----AVYLESDTSCPTTSRSPSGPEKGAHPANSASPPASR 1956
Qy 262 ASVSFLNPNLNCERKEERYEYVPGSTNPEVFKLEDKOPGNMAGNPNLSLQG-CDODA 320
Db 1957 ASLALAGIRQS-----KQPTPG--RADLVSEATDP-----QGIICEKA 1993
Qy 321 QSPGILRLQOVLVQHPQONESNKIYVDLSNERAMSLTIEBRPVKGRKFPVGGFV----376
Db 1994 EKK-----VSDPPQRTNQKIVEISERVPKACGDRPPESDRK--GGFLTQNN 2040
Qy 377 CLESTCSSNLTLTSSGKHKISFLCDDLTRLMNVETISTDHRKYQKSYSLQVP---433
Db 2041 CQEKSAIRLRQSESSPEH-TPEPPSQASQ---VEREI-----RMSFSMAKPATS 2086
Qy 434 -SDILHLPVEL-HDFSK-----LIVPKDRLSLVVPAOKLQOHTHEKCNFSYLV 484
Db 2087 SSSSLQLPALIPESFQGSQMPASVGPKNQVIGL--AGESEPTTPRATRTYS---2141
Qy 485 ASALPSQ-DVYFG-----SFCPGSIRKQIOYKONISVLTRTFADF--XQEASRQGLTV 535
Db 2142 --MPAQFSSHFGEPSHPSPHSPODPQ-----PAMGKLSKXTAKGVT- 2185
Qy 536 SFLPYKEBGFYTTPTKSKVYLRTPNMDRGLPSLTVSW---NISVPDQVACLTF 591
Db 2186 -----NGQGYSVKPLETSKNL-SPVDGRDVSADETSCLIPDKVKTTRQYCCQESW 2238
Qy 592 KERSGVVCGTGRAPMIIOEORTRAEIRFSLDE-----DVLKPSFH-----633
Db 2239 PHES-----TSFSSVKQRIKSFENJANSDRPTAKCATSPFLSVSSKPPINRRSSGSIP 2291
Qy 634 -----HSFWNISNCSPTSGKQDLLFSVTLPRTVDLTVILIAAVGGVLLLSAL 684
Db 2292 SGSPSDMTSRLRSLSGSESGEASGLLPQMTKSSSMTLV-----2335
Qy 685 GLITCYVKKKKKTKNCPA-----VGYNXNINTEMPPROPKFKQGRKNDSHVYA 735
Db 2336 -----SRQNPDDTSNKPSPDPKSLVPVGIPTSTVS--PASPSSKNK-----2376
Qy 736 VIEDTMVYGHLLQDSSGFLQPEVDTPRPGCTMGVCPSPPTIC-----SRAPTA---K 787
Db 2377 -----SSVRAHAPSPVSSKLUOERRTLSPDLDKLNGEDDSASFGAVALFK 2422
Qy 788 LATEBPSPSPSESESEP 805
Db 2423 TQLEITPRRSKGSQATSP 2440

RESULT 9
US-08-755-587-189
; Sequence 189, Application US/08755587
; Patent No. 6045997
; GENERAL INFORMATION:

```

; APPLICANT: Futreal, Phillip A
; APPLICANT: Wooster, Richard F
; APPLICANT: Ashworth, Alan
; APPLICANT: Straton, Michael R
; TITLE OF INVENTION: Materials and methods relating to the
; TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
; NUMBER OF INVENTIONS: susceptibility gene and uses thereof.
; CORRESPONDENCE ADDRESS:
; ADDRESS: Bell Selzer Park & Gibson
; STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,587
; FILING DATE: 25-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9523959.6
; FILING DATE: 23-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525555.0
; FILING DATE: 14-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9617961.9
; FILING DATE: 28-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenneth D Sibley
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5405-135
; INFORMATION FOR SEQ ID NO: 189:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-755-587-189
Query Match          2.6%; Score 114; DB 2; Length 1589;
Best Local Similarity 17.7%; Pred. No. 0.4;
Matches 134; Conservative 104; Mismatches 234; Indels 286; Gaps 34;
QY 198 LPWFHPRNVSGF-----STANRSSIKRLCIHSVEEGSATLMSANYPEGPEDEL 249
DB 558 LPNINEMFEGFCSALGTRKLSVSNELRRKMKLFSDIENSEPRAKVG--PRGFS-- 612
QY 250 MTWQVVAHLRASVSLNPNLSNCRKEER-----VEYIIPGST-----TNPEVKL 297
DB 613 -----AHHDSVAIVFKIKKQNTKESDEKSSKQVTLQNNIEMTTCIFVGRNPEKYIK 664
QY 298 EDKQGNMAGNPNLSIGCCDDQDQSPGILRLQFVLVQHPQESNKIYV---VDLSNER 353
DB 665 NTKHEDSTSSQQRNNLHNSDGSMS-----TSGPVYIHKGDSDL----- 703
QY 354 AMSLITEPRPVQSKRFVPGCFVCLSRTCSSNLTLSGSKHKISFLCDDLTRLMMNVEK 413
DB 704 -----PADQSGK-----CPESCTQYARENTQ-KENIS-----DLTCLMKAE 742
QY 414 TISCTDHRVQQRKYSILOVPSDILHLPEVLDHFSM----- 448
DB 743 T-----CMKSSDKQLPSDKMEQNIKEFNISFOIASGKNTRVSKESLNSKVNIFNRE 794
QY 449 -KLIVPKDRL-SLVLVPAOKLQOHT--HEKPCNTSFSYLVASAIPODLVPGSPCGGS 503
DB 795 TDELTVISDSLNSKILHGINDKMHTSCHK-----ALSIKVFEDH--PIVT 840
QY 504 IKQIQVKNONISVTLRTFAPSFKQZASROGLTVSF-----IPYKE--EGVFTVTPDK 554
DB 841 VSQLPAQGH-----PEYIEISTKEPTLLISFTASGKKVYKINQESLDKYNLFDFTQ 891
```

```

QY 555 SKVYLR--TPNMDRGILSLTSVSNMISVPRDQVACLTFFPKERSGVWCQGRAPMIIQEQRT 613
DB 892 ---YVRKIASPSQGSKPL-----KSKKEELIAYEK-----LEVTAS 925
QY 614 RAEEI---FSIDELVLPKPSFHHHSFWVINSCSPISGK-QLDLFLSVTLTPRTVDTLVI 669
DB 926 KCEMGNFVSKETEMLPQGNHMYRQTEMLTSGTSSKVGEMENIENNVEKNPR----- 978
QY 670 LIAAVGGVLLLSALGLIICCVKKKKKTKNGPVGVIYNNKINTEMPPQPKFGKGR-- 727
DB 979 -----ICCIQSSYPVTDSALAVYTEDSRKTVRE--SLSKGRFWL 1019
QY 728 -----DNDSHVYVI-----EDTM 741
DB 1020 REQGDKLTNTITIKCKEHTEDPAGNASHSLVITETIDTNHVSNOVSTLSDPNV 1079
QY 742 VYGHLLDPS-----SGSFLOPEVDYVRPGTGWGVPSPPTICSRAPPAKLATE 791
DB 1080 CHSVLSQSSFGCHODMHNDSGYFLKNKIDS-----DVPDDM-----KNAEGNTI 1123
QY 792 EPPRSPPESESEPTTFSH-----PNNGDVSSKQTDI 823
DB 1124 SPRVSATKERMLHPQTINVCVQKLETTNTSPHANKDVAI 1161
;
RESULT 10
US-09-538-092-999
; Sequence 999, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glor, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT FILING DATE: 2000-03-29
; PRIOR FILING DATE: 1999-04-10
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-10
; PRIOR APPLICATION NUMBER: 60/178,965
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurateSeqFormatter Version 0.9
; SEQ ID NO 999
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P23515
;
US-09-538-092-999
Query Match          2.6%; Score 112.5; DB 2; Length 440;
Best Local Similarity 19.1%; Pred. No. 0.062;
Matches 79; Conservative 62; Mismatches 144; Indels 129; Gaps 17;
QY 256 VPAHIRASVSFLNF-----NLNCRKEEVE---YIIPGSTNPEV-- 294
DB 49 LPSGIQENIILHNLNSYNHFTDLHNLQTYMTLRLLDISNNLLESLPAHLPRSLNMMSAAN 108
QY 295 ---FKIEDQPGNMAGNPNLSIGCCDD-----AOSPGILRLQFVLVQHPQES 341
DB 109 NNITKLDDK--SDTAYQMLKYLIVSKNMLEKVVLIKNTLRLEVLNLSNKLMTVPTMP 166
QY 342 NKIYVVDISNERAMSLITEPRPVQSKRFVPGCFVCLSRTCSSNLTLSGSKHKISFLC 401
DB 167 SKLHIVDSNN--SLT-----QILPGLTIL-----NLT----- 191
QY 402 DDLTRLMMNVEKTIISCTDHRVQQRKYSILOVPSDILHLPEVLDHFSKVLVPAOKRLSLVL 461
DB 192 -NLTHLYLHNNKFTFIPDSFDQ--LFOIQ-----EITLYNNRWS--CDHKONITVLL 239
```

```

QY 462 VBAQKLOQDTHNEPCNTSSYLVA-SAISODLYPSPCGGGIKOLQVKQINISYLR-- 518
Db 240 KMMETKAVHIGTPCSTQISSLEKHHMYPTSPSGFTSSLFTVSGMQVDIINSLSVTQPK 299
QY 519 -----TPAPSFQXGASRQGLVSPFYP-----FKEE 544
Db 300 VTKIPKQYRKETTKGATLSKDTTFTSTDKAFVPYEDDTSTETINSHEAAATLLIHLQD 359
QY 545 GVFT---VTPDPSKSVYLRTPMMDRGLPSLITSXWMSIYVPDQVACLTPFKHS 595
Db 360 GAVTITSLSSTSKSPPTMILTISGMPN-----NFSENPOOSTLTIINREET 407

```

RESULT 11
US-09-949-016-9282
; Sequence 9282, Application US/09949016

```

; SOFTWARE: FASTSDQ for Windows Version 4.0
; SEQ ID NO 9282
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9282

```

RESULT 12
US-09-560-385A-12

```

: Sequence 12, Application US/09560385A
: Patent No. 6703363
: GENERAL INFORMATION:
: APPLICANT: Boltand, Ariel
: TITLE OF INVENTION: Recombinant Lam
: FILE REFERENCE: 99-274-C
: CURRENT APPLICATION NUMBER: US/09/56
: CURRENT FILING DATE: 2000-04-28
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 12
: LENGTH: 1694
: TYPE: PRT
: ORGANISM: Rattus norvegicus
: US-09-560-385A-12

```

Query Match	2.6%;	Score 112.5;	DB 2;	Length 1694;	
Best Local Similarity	18.8%;	Pred. No. 0.64;	251;	Indels 285;	Gaps 36
Matches	147;	Conservative	99;	Mismatches	251;
QY	128	DVKAHKSIGLELOFSIPRLRQIGBESCPDGVTHSIGRDIATVVRIGTFCSGNGTVSRK	187		
Db	769	DLKGYSISLPLQ---RDLRNGGTED---MFVYIGNKDAKSDYIGMAVVDGLTCVY	822		
QY	188	MOEGVKALHLPMHPHNNVSGFSIANSSTKRLCITESV-----FEG	229		
Db	823	NLGDREAEVQIDOVLTSESSQEAAMDVKCFPRIIQFAKLNTYKATSNKPAPAVYDLEG	882		
QY	230	EGSATYLSAN-----YPEGPEDELTWQVVPAPHLRAS-----VSFLN	268		
Db	883	GSSNTLTNLDEPDAVFYVGYPD-----PELSRLRFPFYKCIELDLNENVLSTYN	936		
QY	269	---FNLNCE---RKEERYVYIPGSTTNEVEFKLEDKQGNAGNFNLSLQGGD	318		
Db	937	FKTTFNINTTEVEECRRRKEESDNGVPEGTO---YARIPYGNAPFPN-----	982		
QY	319	DAOSPLRLQFOV---LVQHPOKESKIYVVDLSNERAM---SLTTEPRVKSRRFP	372		
Db	983	-----IQTQTVYDRGLLFFAEQODNFI-SLTENENLWRYKLNSEPPREKGIIRD--	1032		
QY	373	GCIFYCLESRTCSSNLTTSGSKHKISFLCDDL-TRLMMNV-ECTISCTDHRQCRKSYSL	430		
Db	1033	-----TINDKDSHILITGLQGRMINNERNV-----	1062		
QY	431	QVPSDILHLPEVLEHDFSMKLI--VP---KORSLVLVPAQKLOQHTHE--	473		
Db	1063	-----RIEIEIPDFSTYLLGCIPIAIRERFNISTPAFOCMNLKTKSGVRLNDTVG	1115		
QY	474	--KPCNTSFLYVLSA-----IPSDLL-----FSCFPGSGIKQIYKON	512		
Db	1116	VTKKCSBDMK-LVNTASFSRGGQMSFTNLDDPSTDRQLSFGFTFPQSGTLNHOQR-	1172		
QY	513	ISVTLRTFAPSFQXOASROGLTVSFIYPFKBEVGFVTPDTKSKVYLRTPMDRGLPSLT	572		
Db	1173	TSSLVLVLEBQGHIELSTRD---SNIFIFSPGY-----MD	1205		
QY	573	SVSNNIVPRDQVACLTFFKERSGVQCQTGAEMTIOEQRTAAE---ITSLEDVLKPR	629		
Db	1206	GLLHRVS-----VISDTGRLRLILDQVLRNQRULPSFSMAQOSIRLG	1248		
QY	630	SFHHSFWVNINSNCSPTSGQLDLTFSVTLTPRIVD---TVLLIAVGGVLLSLSLGL	686		
Db	1249	GGHREGCISNV-----LVQFSSQSEVYLDASSYTKDASLGC-----	1286		
QY	687	IICCVKKKKKTKNGPAVGIYNNXNINTEMRQPKKFKGKDKNDSHYAVIEDTMVYGL	746		
Db	1287	--CSL-----NKPRFLMLF-----KSPKRFYKGRIFN-----VNOL	1315		
QY	747	LODSSGFLQPEVDYTPPFQCTMGVCPSPPTICSR-----PTALATEEPPRRSPPE	800		
Db	1316	MODAPQATRSTEA---WODGRSCLPRLNTKASHRALQFGDSPTSHLLKLPOELKLPR	1370		
QY	801	SE 802			

Db 1371 SQ 1372

RESULT 13
US-09-560-385A-10
; Sequence 10, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boudand, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C
; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1725
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-560-385A-10

Query Match 2.6%; Score 112.5; DB 2; Length 1725;
Best Local Similarity 18.8%; Pred. No. 0.66;
Matches 147; Conservative 99; Mismatches 251; Indels 285; Gaps 36;

QY 128 DVKAKSISGELQPSIRPLRQIGPESCPDGVTHSISGRIDATVVRIGTFSGNGTVSRIK 187
DB 800 DLKGTSTLSFLQ--RPDLRENGTED--MFVNYLGNKQASKQYIGMAVVDGQLTCVY 853
QY 188 MOEGVKMALHPWFHPRNVSGFSIARRSSIKRLCIEBSV-----PEG 229
DB 854 NLGDRBAVQIDQVLTSESGEAVMDRVKFORIYFALANTNTKATSNKRPAPAYDLEG 913
QY 230 EGSAITMSAN-----YPEGPEDELMTWQFVPAHLRAS-----VSFLN 268
DB 914 GSSNTLLMLDEBDVAVFYGGYPPD-----FELPSRLRFPKPGCIEILDINENVLSLYN 967
QY 269 -----FNLSNCE-----RKEREVEYIIGSTTNPEVFLDEKQPMNMGNFNLSQSGDQ 318
DB 968 FKTTFNMTTVEPCRRKESDKNYFEGTG-----VARIPTQPAAPPNF----- 1013
QY 319 DAQSPGILRLQFOV---LVQHPONESNKIYVVDLSNERAM---SLTIEPRPVKQSRKEVP 372
DB 1014 -----IOTIQTVDROGLLFFAENQDNFI-SLNIEDGNLMVRKYKLNSEPPKEKIRD--- 1063
QY 373 GCFVCLSESRTCSSNLTITSGSKHKISFLCDDL-TRLMANV-EKTICTDHRYCQKXSYSL 430
DB 1064 -----TINDGKXHSILITIGKLOKMMINVNERSV----- 1093
QY 431 QVPSDILHLPEVLIHDSMKL--VP---KDRLSLVLPQAQKQOHTHE----- 473
DB 1094 -----RIEGIFPFSYIYAGIPIAIRERNISTPAPQGCMLKKTSGVRLNDTVG 1146
QY 474 --KPCNTSPSYLVASA-----IPSDLY-----FGSFCGSGSIKQIQVKON 512
DB 1147 VTKKCSBDMK-LVTRASFSRGGQMSFTMLDVPSTDRFQLSRFGQRFQSGTILAHQTR-- 1203
QY 513 ISVTLRTPASFXQEASRQGLTVSFIPIYFKBEQVTVTPDTKSKYVLTPTKMDRGLPSLT 572
DB 1204 TSSLVLTLEDGHIESTRD---SNIPFKSPGT-----MD 1236
QY 573 SVSNNISVPRDQVALTFPKERSGVVCOGNRAFMIIQORTRAE--IFSLDEVDLKP 629
DB 1237 GLLHNVS-----VISDTSGRLILIDQVLRNRRLPSFSNAQOSLRG 1279
QY 630 SFHHHSFVNINSCSPTSKQDLDFSVTLTPRTVDL---TVILIAVGGVLLLSALGL 686
DB 1280 GGHFEGCISNV-----LVQRFSGSEVLDLASKTKKXASLGG----- 1317
QY 687 IICCYKAKKKKTKNGPANGIYKXNINTEMPROPKFQKGRKNDSHVAVIEDTWYGH 746
DB 1318 --CSL-----NKBPFLMLF-----KSPKRFNKGRIFN-----VNO 1346

QY 747 LODSSGSLQPEVDTPRPOGTMGVCPSPPTICSA-----PTAKLATEPPRSPPE 800
DB 1347 WQDAPQATRSTEA-----WQDGRSCLPPLNTKASHRALQFDSFTSHLLKLPQELKPR 1401

QY 801 SE 802
DB 1402 SQ 1403

RESULT 14
US-09-949-016-6138
; Sequence 6138, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6138
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6138

Query Match 2.5%; Score 108; DB 2; Length 699;
Best Local Similarity 18.3%; Pred. No. 0.39;
Matches 82; Conservative 58; Mismatches 145; Indels 164; Gaps 22;

QY 235 LMSANYPGPEDELMTWQFVPAHLRASVSFLNPNLSN---CERKEEREVEYIIGSTTN 291
DB 31 IQSPGYPPSPYSDBEVNTNITVPDGFRIKLYFMHFNLSYLCEDYVKE----- 81
QY 292 PEVFLDEKQCGMNAAGNLSLQGCCDDQDQSPGILRLQFQVLVQHPONESNKIYVVDLSN 351
DB 82 -----TEQVATATCGRTTDTT-----QIRG-----QEVVLSFGSFMSTIRSDPSN 124
QY 352 ERA-----MSLTIEPRPVKQSRKF-----VPGCFVC-----LESRTCSN 386
DB 125 EERFTGFDAHYMAVDVDECKERDEBELSCDHYCHNYIGGYCSCRFGYLHTDMRTCC-- 181
QY 387 LTLTSGSKHKISFLCDDLTRLMANVVEKTICTDHRYCQKXSYSLQVPSDILHLPELHDF 446
DB 182 -----RVECSDMLFQIR-----TGVTISP---DF 202
QY 447 SMKLLVPKORSLVLPQAK-----LQOHTHERPCNTSPSYLVASAPSD 493
DB 203 PNP--YKSSSECLYIIEBEGFMNLOREFIDFDQDH-EVPC--PYVIKIKYGPK--- 254
QY 494 YFGSFCPGSGSIKQIQVKONISVTLRTFAPSFXQEASRQGLTVSF----- 537
DB 255 VLGFGCGEKAPEPIST-QSHSVLLILFHSN---SAENRGWRLSYRAAGNECPDELQPVHG 310
QY 538 -----IPYFKBEQVTVTPDTKSKYVLTPTKMNMR-GLPSLTSVSNISVPRDQVA-C-- 587
DB 311 KIEPSQAKYFPKQDVL-VSCDTGYKVLKDNVEMPTFOIECLKDGTWSKXKIFPTCKIVDCRA 369

QY 588 -----LTFPKERSGVVQ 600
DB 370 PGELEHGLITSTNNLITTYKSELKYSQ 398

RESULT 15

US-09-248-796A-20641
; Sequence 20641, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANUM
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20641
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20641

Query Match 2.5%; Score 107.5; DB 2; Length 649;
Best Local Similarity 17.9%; Pred. No. 0.38; Mismatches 206; Indels 185; Gaps 22;
Matches 102; Conservative 77;

Dy 348 DLS--NERAMSLTEBPVKQS-RKVFQCFVCL-----ESRSCSNLT 389
|||:::|||||:::|||||:::|||||:::|||||
Db DLSAVLDQQGQCSEYENPPTTOLQIKIRKGLVDLQKILYMAABQEATDSTNL 153
|||:::|||||:::|||||:::|||||:::|||||
Dy 390 TSGSKHKTSFLCDLLTRLMANVEKTISCTDHRYC-----ORKSYSLQVPSDILLHP 440
|||:::|||||:::|||||:::|||||:::|||||
Db 154 YG-----ALIQDDLNMLKDNNNISNGSNGSRFALKSEIETNGHQXSEKQTNDKDNOD 207
|||:::|||||:::|||||:::|||||:::|||||
Dy 441 VELHD-----FSWKLIVPKDRSLVLVPKQLQH----- 470
|||:::|||||:::|||||:::|||||:::|||||
Db 208 IEMTDALNKVEQOGEIEFFETQLSPQ---LKRAPSLTCDWHPTTEVFAYGKDGDTATI 263
|||:::|||||:::|||||:::|||||:::|||||
Dy 471 -----THEKPCTSFYSYLASAIPQDLY-----FGSFCPGSGSIKOIQ 508
|||:::|||||:::|||||:::|||||:::|||||
Db 264 NAIDQKIERTLTRPLNLINKQINIVSWSPGSLITTCGANSELRAWSPDKLR--- 320
|||:::|||||:::|||||:::|||||:::|||||
Dy 509 VKONISVTLTFAPSFOEASRGQLVVSIFPYKEEGVFTVTPDKSKVYLRTPNMDRGL 568
|||:::|||||:::|||||:::|||||:::|||||
Db 321 ---NIASITVDEAISL--ENTKLISISSLWS-SPTGKFLSIDSLNOVCI---WDGTT 370
|||:::|||||:::|||||:::|||||:::|||||
Dy 569 PSULTSVSNISVPRDVACLTFFKE-RSGVICGTG--RAFMIIOEQTRAELFSLDEDV 625
|||:::|||||:::|||||:::|||||:::|||||
Db 371 ISLKKQIKMILEINDSVVSCMLSBDKFAVTTNTNGIKIYDILOPPSSH----FGNQLDV 426
|||:::|||||:::|||||:::|||||:::|||||
Dy 626 LKPSFHHSFWNVINSNCSPTSKGQDLDFSVTLTPRVDTLVLLIAVGCGVLLSALG 685
|||:::|||||:::|||||:::|||||:::|||||
Db 427 OPTGLNGHEHNISIMKLNIPVT-----KLATCSDPDYSI---KWSSSSSQECLED 474
|||:::|||||:::|||||:::|||||:::|||||
Dy 686 LIICVVKKKKTKTGPAVGIVY-----NXNI-NTEM 715
|||:::|||||:::|||||:::|||||:::|||||
Db 475 LNINPEKKYTLKHAAPMIGLIWL.PDSNSNRSLLSVSMGIINTDAKTSNIKRSSEL 534
|||:::|||||:::|||||:::|||||:::|||||
Dy 716 PRQPKFKQGRKNDSDSHVAVIEDTMYYGHLLQDSSGFL----- 755
|||:::|||||:::|||||:::|||||:::|||||
Db 535 FNKNKDNFTNELKDDNNNDTIKIDVLVFNNAV-L.SPDGYLALGDYCRVTIMDVDTTHYL 593
|||:::|||||:::|||||:::|||||:::|||||
Dy 756 -----QPEVDTYRPFGGTMGVC 772
|||:::|||||:::|||||:::|||||:::|||||
Db 594 NEPKDLIRCKAAVYRKPEISAEDKLKATIGIC 623

Search completed: May 4, 2006, 22:27:24
Job time : 50 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 22:27:41 ; Search time 84 Seconds
(without alignments)
4158.397 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4377
Sequence: 1 MAGNCGVSIALGVLLGA.....SSKOTDIFLXTQEPMEPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.dep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.dep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.dep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.dep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.dep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.dep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4375	99.9	836	5	US-10-781-564-1 Sequence 1, Appli
2	4374	99.9	836	3	US-09-899-569A-4 Sequence 4, Appli
3	3625.5	82.8	709	4	US-10-351-334-132 Sequence 132, App
4	3454	78.9	749	3	US-09-899-569A-2 Sequence 2, Appli
5	2155	49.2	414	4	US-10-351-334-273 Sequence 273, App
6	2155	49.2	443	4	US-10-351-334-271 Sequence 271, App
7	1779	40.6	343	4	US-10-028-072-162 Sequence 162, App
8	1779	40.6	343	4	US-10-140-808-162 Sequence 162, App
9	1779	40.6	343	4	US-10-121-049-162 Sequence 162, App
10	1779	40.6	343	4	US-10-123-904-162 Sequence 162, App
11	1779	40.6	343	4	US-10-140-470-162 Sequence 162, App
12	1779	40.6	343	4	US-10-175-746-162 Sequence 162, App
13	1779	40.6	343	4	US-10-176-918-162 Sequence 162, App
14	1779	40.6	343	4	US-10-176-921-162 Sequence 162, App
15	1779	40.6	343	4	US-10-137-865-162 Sequence 162, App
16	1779	40.6	343	4	US-10-140-474-162 Sequence 162, App
17	1779	40.6	343	4	US-10-142-431-162 Sequence 162, App
18	1779	40.6	343	4	US-10-143-114-162 Sequence 162, App
19	1779	40.6	343	4	US-10-142-419-162 Sequence 162, App
20	1779	40.6	343	4	US-10-123-262-162 Sequence 162, App
21	1779	40.6	343	4	US-10-142-423-162 Sequence 162, App
22	1779	40.6	343	4	US-10-121-050-162 Sequence 162, App
23	1779	40.6	343	4	US-10-141-755-162 Sequence 162, App
24	1779	40.6	343	4	US-10-143-032-162 Sequence 162, App
25	1779	40.6	343	4	US-10-123-108-162 Sequence 162, App
26	1779	40.6	343	4	US-10-123-236-162 Sequence 162, App
27	1779	40.6	343	4	US-10-123-261-162 Sequence 162, App

28	1779	40.6	343	4	US-10-140-921-162 Sequence 162, App
29	1779	40.6	343	4	US-10-140-928-162 Sequence 162, App
30	1779	40.6	343	4	US-10-121-045-162 Sequence 162, App
31	1779	40.6	343	4	US-10-123-292-162 Sequence 162, App
32	1779	40.6	343	4	US-10-123-903-162 Sequence 162, App
33	1779	40.6	343	4	US-10-124-819-162 Sequence 162, App
34	1779	40.6	343	4	US-10-124-822-162 Sequence 162, App
35	1779	40.6	343	4	US-10-140-925-162 Sequence 162, App
36	1779	40.6	343	4	US-10-160-458-162 Sequence 162, App
37	1779	40.6	343	4	US-10-124-824-162 Sequence 162, App
38	1779	40.6	343	4	US-10-127-825A-162 Sequence 162, App
39	1779	40.6	343	4	US-10-127-829A-162 Sequence 162, App
40	1779	40.6	343	4	US-10-127-835A-162 Sequence 162, App
41	1779	40.6	343	4	US-10-127-839A-162 Sequence 162, App
42	1779	40.6	343	4	US-10-127-901A-162 Sequence 162, App
43	1779	40.6	343	4	US-10-128-633A-162 Sequence 162, App
44	1779	40.6	343	4	US-10-131-813A-162 Sequence 162, App
45	1779	40.6	343	4	US-10-131-818A-162 Sequence 162, App

ALIGNMENTS

```
RESULT 1
US-10-781-564-1
; Sequence 1, Application US/10781564
; Publication No. US20040247601A1
; GENERAL INFORMATION:
; APPLICANT: QuiGley, James P.
; APPLICANT: Hooper, John D.
; APPLICANT: Testa, Jacqueline E.
; TITLE OF INVENTION: The Scripps Research Institute
; FILE REFERENCE: 1361.036US1
; CURRENT APPLICATION NUMBER: US/10/781,564
; PRIOR FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 60/448,828
; PRIOR FILING DATE: 2003-02-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-781-564-1

Query Match      99.9% Score 4375; DB 5; Length 836;
Beet Local Similarity 99.6%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGVLLGAARLPRGAEPRIALPRESNITVLTKLGTPTLLAKPCYVI 60
    |||||
DB 1 MAGNCGVSIALGVLLGAARLPRGAEPRIALPRESNITVLTKLGTPTLLAKPCYVI 60
    |||||

QY 61 SKRHITMISGGERIVTFSCQSPENHFVIEIKNDICMGCPFPFGVOLOPSTSLPT 120
    |||||
DB 61 SKRHITMISGGERIVTFSCQSPENHFVIEIKNDICMGCPFPFGVOLOPSTSLPT 120
    |||||

QY 121 LNRFTIIVYKHKHSIGLELQPSIRPLRQIGGSCPDGVTHSISGRIDATVVRIGTFCSN 180
    |||||
DB 121 LNRFTIIVYKHKHSIGLELQPSIRPLRQIGGSCPDGVTHSISGRIDATVVRIGTFCSN 180
    |||||

QY 181 GTVSRIKMQEGVKALHLPMFHPRVNVSFISIANSSIKRLCIISVFEGESATLMSANY 240
    |||||
DB 181 GTVSRIKMQEGVKALHLPMFHPRVNVSFISIANSSIKRLCIISVFEGESATLMSANY 240
    |||||

QY 241 PEGPFEDLMTWQFVPHLAASVSLFNFNLSNCRKEERYIIPGTTNPEVKLEDK 300
    |||||
DB 241 PEGPFEDLMTWQFVPHLAASVSLFNFNLSNCRKEERYIIPGTTNPEVKLEDK 300
    |||||

QY 301 QPGNMGAFNLSLOGCDODDASPGILRLQFOVLVOHPONESNKIYVVDLSNRAMSLTIE 360
    |||||
DB 301 QPGNMGAFNLSLOGCDODDASPGILRLQFOVLVOHPONESNKIYVVDLSNRAMSLTIE 360
    |||||
```

QY 361 PRPVQSRKRVPGCVCLSESRCTSSNLTLTSGSKHKISFLCDLRLMNMNVEKTSCTDH 420
DB 361 PRPVQSRKRVPGCVCLSESRCTSSNLTLTSGSKHKISFLCDLRLMNMNVEKTSCTDH 420
QY 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPKORLSLVLPQAOKLQOHTHEKPCNTSF 480
DB 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPKORLSLVLPQAOKLQOHTHEKPCNTSF 480
QY 481 SYLVASAIPIPSODLYFGSFCPGGSIKOIQVKONI SVTLRTFAPSFKOASROGLTVSFIPY 540
DB 481 SYLVASAIPIPSODLYFGSFCPGGSIKOIQVKONI SVTLRTFAPSFKOASROGLTVSFIPY 540
QY 541 FKEEGVFVTPDTSKAVYLRTPNMWDRGLPSLTSVSNISVPRDQVACLFFKERSGVQC 600
DB 541 FKEEGVFVTPDTSKAVYLRTPNMWDRGLPSLTSVSNISVPRDQVACLFFKERSGVQC 600
QY 601 TGRAPMIIOEORTRAEIEFSLDEVDLPRKPSFHHHSFWVNISNCSPTSGQDLLEFSVLT 660
DB 601 TGRAPMIIOEORTRAEIEFSLDEVDLPRKPSFHHHSFWVNISNCSPTSGQDLLEFSVLT 660
QY 661 PRTVDTLVLLIAVGGVLLLSALGLIICVKKKKKKTKNGPAVGITYNKNINTEMPROK 720
DB 661 PRTVDTLVLLIAVGGVLLLSALGLIICVKKKKKKTKNGPAVGITYNKNINTEMPROK 720
QY 721 KFOGKNDKNDSHVYAVIEDTWYGHLLQDSSGSLQPEVDYTRPPGTMGVCPSPPTIC 780
DB 721 KFOGKNDKNDSHVYAVIEDTWYGHLLQDSSGSLQPEVDYTRPPGTMGVCPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPSESESEPTFSHPNNGDVSSKDTDIPLXTOEPMEPAE 836
DB 781 SRAPTAKLATEEPPRSPSESESEPTFSHPNNGDVSSKDTDIPLXTOEPMEPAE 836

RESULT 2
US-09-899-569A-4
; Sequence 4, Application US/09899569A
; Patent No. US20020142003A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020142003Albert Schweizer
; APPLICANT: Marwa Scherl-Mostagheer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Absenher
; TITLE OF INVENTION: Tumorssoziiertes Antigen (B345)
; FILE REFERENCE: 0652.2280001
; CURRENT APPLICATION NUMBER: US/09/899,569A
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: DE 100 33 080. 0
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 101 19 294. 0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/297,747
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-899-569A-4

Query Match 99.9%; Score 4374; DB 3; Length 836;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAGLNCGVATILGVLILGAARLRGAAFEIALPRESNITVYIKGTPTLAKPCYIY 60
DB 1 MAGLNCGVATILGVLILGAARLRGAAFEIALPRESNITVYIKGTPTLAKPCYIY 60
QY 61 SKRHITMLSIKSGERIVFTFSCSPENHVIIEIOKNIDCMGCPGEGEVOLOPSTILPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCSPENHVIIEIOKNIDCMGCPGEGEVOLOPSTILPT 120

DB 61 SKRHITMLSIKSGERIVFTFSCSPENHVIIEIOKNIDCMGCPGEGEVOLOPSTILPT 120
QY 121 LNRTITMVKAKHKSIGLELOFSIPRLROIIGCEGSPDVTMISIRIATATVRICTPCSN 180
DB 121 LNRTITMVKAKHKSIGLELOFSIPRLROIIGCEGSPDVTMISIRIATATVRICTPCSN 180
QY 181 GTVSRIKMOEGVKMALHLPMFHPRVNVSFGSILANRSIRKLCTIEVSFEGEGSATLMSANY 240
DB 181 GTVSRIKMOEGVKMALHLPMFHPRVNVSFGSILANRSIRKLCTIEVSFEGEGSATLMSANY 240
QY 241 PEGPEDELMTWQFVVPALRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLEDK 300
DB 241 PEGPEDELMTWQFVVPALRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLEDK 300
QY 301 QPGNNAGNPNLSLOCCDDAOSPGLRLQFOLVOPHONESKITYVDLSNRAASLTIE 360
DB 301 QPGNNAGNPNLSLOCCDDAOSPGLRLQFOLVOPHONESKITYVDLSNRAASLTIE 360
QY 361 PRPVQSRKRVPGCVCLSESRCTSSNLTLTSGSKHKISFLCDLRLMNMNVEKTSCTDH 420
DB 361 PRPVQSRKRVPGCVCLSESRCTSSNLTLTSGSKHKISFLCDLRLMNMNVEKTSCTDH 420
QY 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPKORLSLVLPQAOKLQOHTHEKPCNTSF 480
DB 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPKORLSLVLPQAOKLQOHTHEKPCNTSF 480
QY 481 SYLVASAIPIPSODLYFGSFCPGGSIKOIQVKONI SVTLRTFAPSFKOASROGLTVSFIPY 540
DB 481 SYLVASAIPIPSODLYFGSFCPGGSIKOIQVKONI SVTLRTFAPSFKOASROGLTVSFIPY 540
QY 541 FKEEGVFVTPDTSKAVYLRTPNMWDRGLPSLTSVSNISVPRDQVACLFFKERSGVQC 600
DB 541 FKEEGVFVTPDTSKAVYLRTPNMWDRGLPSLTSVSNISVPRDQVACLFFKERSGVQC 600
QY 601 TGRAPMIIOEORTRAEIEFSLDEVDLPRKPSFHHHSFWVNISNCSPTSGQDLLEFSVLT 660
DB 601 TGRAPMIIOEORTRAEIEFSLDEVDLPRKPSFHHHSFWVNISNCSPTSGQDLLEFSVLT 660
QY 661 PRTVDTLVLLIAVGGVLLLSALGLIICVKKKKKKTKNGPAVGITYNKNINTEMPROK 720
DB 661 PRTVDTLVLLIAVGGVLLLSALGLIICVKKKKKKTKNGPAVGITYNKNINTEMPROK 720
QY 721 KFOGKNDKNDSHVYAVIEDTWYGHLLQDSSGSLQPEVDYTRPPGTMGVCPSPPTIC 780
DB 721 KFOGKNDKNDSHVYAVIEDTWYGHLLQDSSGSLQPEVDYTRPPGTMGVCPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPSESESEPTFSHPNNGDVSSKDTDIPLXTOEPMEPAE 836
DB 781 SRAPTAKLATEEPPRSPSESESEPTFSHPNNGDVSSKDTDIPLXTOEPMEPAE 836

RESULT 3
US-10-351-334-132
; Sequence 132, Application US/10351334
; Publication No. US20040034196A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P2
; CURRENT APPLICATION NUMBER: US/10/351,334
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12

PRIOR APPLICATION NUMBER: 60/095,454
PRIOR FILING DATE: 1998-08-06
PRIOR APPLICATION NUMBER: 60/095,455
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 132
LENGTH: 709
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (189)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (275)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (414)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (438)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (641)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (643)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (696)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (697)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-351-334-132
Query Match 82.8%; Score 3625.5; DB 4; Length 709;
Best Local Similarity 98.3%; Pred. No. 5.5e-315;
Matches 698; Conservative 1; Mismatches 10; Indels 1; Gaps 1;
QY 1 MAGINCGVSIALLGLVLLGARLPARGAFAFIALPRESNITVLTGTPTLAKCYVI 60
DB 1 MAGINCGVSIALLGLVLLGARLPARGAFAFIALPRESNITVLTGTPTLAKCYVI 60
QY 61 SKRHTMTLSKSGERIVTFSCSPENHFVIEIQNIDCMGSPCFGEVOLQPSLPT 120
DB 61 SKRHTMTLSKSGERIVTFSCSPENHFVIEIQNIDCMGSPCFGEVOLQPSLPT 120
QY 121 LNRFTFMDVKAHKSIGLEQFSIPRLRQIGPGECPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTFMDVKAHKSIGLEQFSIPRLRQIGPGECPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVARIKQEGVKALHPMFHPRNVSGFSIARNSIKRLCIESEVEGEGSATLMSANY 240
DB 181 GTVARIKQEGVKALHPMFHPRNVSGFSIARNSIKRLCIESEVEGEGSATLMSANY 240
QY 241 PEGFEDELMTQFVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLEDK 300
DB 241 PEGFEDELMTQFVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLEDK 300
QY 301 QPQNMAGFNLSLQCCDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTE 360
DB 301 QPQNMAGFNLSLQCCDDAOSPGLRLQFOVLVQHPONESNKIYVVDLSNERAMSLTE 360
QY 361 PRVYQSKRFVPGCFVCLSERTCSNLTLTSGSKHKISFLDDDLRLMNMVETKISCTDH 420

DB 361 PRVYQSKRFVPGCFVCLSERTCSNLTLTSGSKHKISFLDDDLRLMNMVETKISCTDH 420
QY 421 RYQQRKSYSLQVPSDILHLPEVLHDFSWKLLVPRXDRSLVLPVQKLOOHTHEKPCNTSF 480
DB 421 RYQQRKSYSLQVPSDILHLPEVLHDFSWKLLVPRXDRSLVLPVQKLOOHTHEKPCNTSF 480
QY 481 SYLVASAIPSQDLYFGSFCPGGSIKQIQVKONISVTLTFTFAPSPFOEASRQGLTVSFI 540
DB 481 SYLVASAIPSQDLYFGSFCPGGSIKQIQVKONISVTLTFTFAPSPFOEASRQGLTVSFI 540
QY 541 FKREGVFTVPTDTSKYVLRTPNMDRGLPSLTSSVSNISVPRDOVACLTFFKESGVYQ 600
DB 541 FKREGVFTVPTDTSKYVLRTPNMDRGLPSLTSSVSNISVPRDOVACLTFFKESGVYQ 600
QY 601 TGRAFMIIQEORTAEIEFSLDEBVLPKPSFHHHSFWNINXSPGTSKQDLDFSVTLT 660
DB 601 TGRAFMIIQEORTAEIEFSLDEBVLPKPSFHHHSFWNINXSPGTSKQDLDFSVTLT 660
QY 661 PRTVDLTVILIAAVGGVLLSALGLIICVYKXKXKTKNCPAVGIYXN 710
DB 661 PRTVDLTVILIAAVGGVLLSALGLIICVYKXKXKTKNCPAVGIYXN 710

RESULT 4
US-09-899-569A-2

Sequence 2, Application US/09899569A
Patent No. US20020142003A1
GENERAL INFORMATION:
APPLICANT: No. US20020142003Albert Schweitzer
APPLICANT: Marwa Scherl-Mostagaeer
APPLICANT: Wolfgang Sommergruber
APPLICANT: Roger Abseher
TITLE OF INVENTION: Tumorcassozieretes Antigen (B345)
FILE REFERENCE: 0652,2280001
CURRENT APPLICATION NUMBER: US/09/899,569A
PRIOR FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: DE 100 33 080.0
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: DE 101 19 294.0
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 60/243,158
PRIOR FILING DATE: 2000-10-25
PRIOR APPLICATION NUMBER: US 60/297,747
PRIOR FILING DATE: 2001-06-14
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 749
TYPE: PRT
ORGANISM: Homo sapiens
US-09-899-569A-2

Query Match 78.9%; Score 3454; DB 3; Length 749;
Best Local Similarity 90.0%; Pred. No. 1.4e-299;
Matches 674; Conservative 9; Mismatches 30; Indels 36; Gaps 6;
QY 67 MLSTKSGERIVTFSCSPENHFVIEIQNIDCMGSPCFGEVOLQPSLPTLNRFTFI 126
DB 1 MLSTKSGERIVTFSCSPENHFVIEIQNIDCMGSPCFGEVOLQPSLPTLNRFTFI 126
QY 127 WDVAKAHKSIGLEQFSIPRLRQIGPGECPDGVTHSISGRIDATVVRIGTFCSNGVRSI 186
DB 127 WDVAKAHKSIGLEQFSIPRLRQIGPGECPDGVTHSISGRIDATVVRIGTFCSNGVRSI 186
QY 187 KMOEGVGMALHPMFHPRNVSGFSIARNSIKRLCIESEVEGEGSATLMSANYPEGFPE 246
DB 187 KMOEGVGMALHPMFHPRNVSGFSIARNSIKRLCIESEVEGEGSATLMSANYPEGFPE 246
QY 247 DELMTQFVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLEDKQPGNMA 306
DB 247 DELMTQFVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLEDKQPGNMA 306
QY 306 DELMTQFVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLEDKQPGNMA 360
DB 306 DELMTQFVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLEDKQPGNMA 360

Qy	30	GNFNLSLOGDDOAGSPGLILQFOVLYOHOPNESNKIYYVDLSNERMSLTIIEPRVYK	366
Db	241	GNFNLSLOGDDOAGSPGLILQFOVLYOHOPNESNKIYYVDLSNERMSLTIIEPRVYK	3000
Qy	367	SRKFVPGCFVGLIESRTCCSNLTVTSGSKKISFLCDDLTRLMNVYEKTIISCTDHRVYQK	426
Db	301	SRKFVPGCFVGLIESRTCCSNLTVTSGSKKISFLCDDLTRLMNVYEKTIISCTDHRVYQK	3600
Qy	427	SYSLQVPSDILHLPEVLEHPSWMLVAPKRLSLVVPPOKLOQHHEKPCNTSHSYVAS	466
Db	361	SYSLQVPSDILHLPEVLEHPSWMLVAPKRLSLVVPPOKLOQHHEKPCNTSHSYVAS	4220
Qy	487	ALPSODLYFGSSPCPGSIIKOIQVKONI SVTLRTFAPSPXQEASRGLTVSFI PYFKEGV	546
Db	421	ALPSODLYFGSSPCPGSIIKOIQVKONISVTLRTFAPSPXQEASRGLTVSFI PYFKEGV	480
Qy	547	FTVTPDTRKSUVYLRTNMMRGVLSLTVSFWNISVPRDVACLTFPKHSGVVCOTGRFAM	606
Db	481	FTVTPDTRKSUVYLRTNMMRGVLSLTVSFWNISVPRDVACLTFPKHSGVVCOTGRFAM	5400
Qy	607	IIQORTRAEIEFLSDEDEVLPKPSFFHHSFWNINSNCSPTSGKODLFLSVTLTPRTVDL	666
Db	541	IIQORTRAEIEFLSDEDEVLPKPSFFHHSFWNINSNCSPTSGKODLFLSVTLTPRTVDL	6000
Qy	667	TVLIIAAGVGVLIIISALGLIICQVKKKKKTKNKPAAGIYNXNINTEMPRQPKKFOGR	726
Db	601	TVLIIAAGVGVLIIISALGLIICQVKKKKKTKNKPAAGIYNXNINTEMPRQPKKFOGR	655
Qy	727	KDNDSHYAVIIEPTMYGHLLODSGSGFLQPEVDTYRPFQGTMC---PSP-----	776
Db	660	-----ERTMT--PFCMQSSRT-----PWY--MGICVRI PAAPSCSQRT	694
Qy	777	PTICSRAPTKATIEEPPRSPRESSEP	805
Db	695	PTICSRAPMGWSVLPPHPVAPGFOQLSWP	723

PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4,le-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGLVLLGAAFLPGAGAPFIALPRESNITVLLKGTPTLLAKPCYVI 60
DB 1 MAGLNCGVSIALLGLVLLGAAFLPGAGAPFIALPRESNITVLLKGTPTLLAKPCYVI 60
QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQNIDCMGQPCFGEVQLQPSLTP 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQNIDCMGQPCFGEVQLQPSLTP 120
QY 121 LNRFTIWDVKAKHSIGLELQPSIPRLQIGCESCPDCVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWDVKAKHSIGLELQPSIPRLQIGCESCPDCVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPMFHPRVVSGFSIANKSSIRLCLIESVFEGGSATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPMFHPRVVSGFSIANKSSIRLCLIESVFEGGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCKEERVEYYIPGSTNPEVKLEBK 300
DB 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCKEERVEYYIPGSTNPEVKLEBK 300
QY 301 QPGMAGNFNLSLQCCDDAOSPGTLRQFOVLVQHPQNSNK 343
DB 301 QPGMAGNFNLSLQCCDDAOSPGTLRQFOVLVQHPQNSNK 343

RESULT 8

US-10-140-808-162

Sequence 162, Application US/10140808

Publication No. US20030017563A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tunes, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3330R1C182

CURRENT APPLICATION NUMBER: US/10/140,808

CURRENT FILING DATE: 2002-05-07

Prior Application removed - See file wrapper or Palm

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 162

LENGTH: 343

TYPE: PRT

ORGANISM: Homo Sapien

US-10-140-808-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4.1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPARGAEAFIALPRESNTIVLILKGTPTLLAKPCYIV 60
DB 1 MAGNCGVSIALGLVLLGAARLPARGAEAFIALPRESNTIVLILKGTPTLLAKPCYIV 60

QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120

QY 121 LNRFTIWDVKAHKSIGLELQFSIPRLRQIGGESCPCDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWDVKAHKSIGLELQFSIPRLRQIGGESCPCDGVTHSISGRIDATVVRIGTFCSN 180

QY 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIRLCTIESVFEGEGSATLMSANY 240

QY 241 PEGPPEDELMTWQFVVPALRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLBDK 300
DB 241 PEGPPEDELMTWQFVVPALRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLBDK 300

QY 301 QPGMAGNFNLSLQCCDDAOSPGLRLQFOVLVQHNPONESK 343
DB 301 QPGMAGNFNLSLQCCDDAOSPGLRLQFOVLVQHNPONESSE 343

RESULT 9
US-10-121-049-162
; Sequence 162; Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4.1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPARGAEAFIALPRESNTIVLILKGTPTLLAKPCYIV 60
DB 1 MAGNCGVSIALGLVLLGAARLPARGAEAFIALPRESNTIVLILKGTPTLLAKPCYIV 60

QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120

DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120

QY 121 LNRFTIWDVKAHKSIGLELQFSIPRLRQIGGESCPCDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWDVKAHKSIGLELQFSIPRLRQIGGESCPCDGVTHSISGRIDATVVRIGTFCSN 180

QY 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIRLCTIESVFEGEGSATLMSANY 240

QY 241 PEGPPEDELMTWQFVVPALRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLBDK 300
DB 241 PEGPPEDELMTWQFVVPALRASVSFLNPNLSNCRKEERVEYYIPGSTTNPVEFKLBDK 300

QY 301 QPGMAGNFNLSLQCCDDAOSPGLRLQFOVLVQHNPONESK 343
DB 301 QPGMAGNFNLSLQCCDDAOSPGLRLQFOVLVQHNPONESSE 343

RESULT 10
US-10-123-904-162
; Sequence 162; Application US/10123904
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-904-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4.1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPARGAEAFIALPRESNTIVLILKGTPTLLAKPCYIV 60
DB 1 MAGNCGVSIALGLVLLGAARLPARGAEAFIALPRESNTIVLILKGTPTLLAKPCYIV 60

QY 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGSPCFGEVQLQPSSTLLPT 120

QY 121 LNRFTIWDVKAHKSIGLELQFSIPRLRQIGGESCPCDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWDVKAHKSIGLELQFSIPRLRQIGGESCPCDGVTHSISGRIDATVVRIGTFCSN 180

QY 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPMFHPRNVSQFSIANRSSIRLCTIESVFEGEGSATLMSANY 240

```
Db      181  GTVSRIKMOEGVKMALHLPMFHPNRVSGFSIANRSSIKRLCIIESVFEGSGSATLMSANY 240
Qy      241  PEGFPEDELMTWQFVVPAPHLRASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
Db      241  PEGFPEDELMTWQFVVPAPHLRASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
Qy      301  QPGNMGNFNLSLQGCDDAOSPGLRLQFOVLVOHPONESNK 343
Db      301  QPGNMGNFNLSLQGCDDAOSPGLRLQFOVLVOHPONESSE 343

RESULT 11
US-10-140-470-162
; Sequence 162, Application US/10140470
; Publication No. US2003002231A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Defoige, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-470-162

Query Match      40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4,1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAGNCGVSIALGLVLLGAARLPGABAEFIALPRESNITVLIKLGPTLLAKPCYIV 60
Db      1  MAGNCGVSIALGLVLLGAARLPGABAEFIALPRESNITVLIKLGPTLLAKPCYIV 60
Qy      61  SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGCPFGEVQLQPSLSLPT 120
Db      61  SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGCPFGEVQLQPSLSLPT 120
Qy      121  LNRFTIWDVKAKHSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db      121  LNRFTIWDVKAKHSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Qy      121  LNRFTIWDVKAKHSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db      121  LNRFTIWDVKAKHSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Qy      181  GTVSRIKMOEGVKMALHLPMFHPNRVSGFSIANRSSIKRLCIIESVFEGSGSATLMSANY 240
Db      181  GTVSRIKMOEGVKMALHLPMFHPNRVSGFSIANRSSIKRLCIIESVFEGSGSATLMSANY 240
Qy      241  PEGFPEDELMTWQFVVPAPHLRASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
Db      241  PEGFPEDELMTWQFVVPAPHLRASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
Qy      301  QPGNMGNFNLSLQGCDDAOSPGLRLQFOVLVOHPONESNK 343
Db      301  QPGNMGNFNLSLQGCDDAOSPGLRLQFOVLVOHPONESSE 343
```

```
RESULT 12
US-10-175-746-162
; Sequence 162, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Defoige, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-746-162

Query Match      40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4,1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAGNCGVSIALGLVLLGAARLPGABAEFIALPRESNITVLIKLGPTLLAKPCYIV 60
Db      1  MAGNCGVSIALGLVLLGAARLPGABAEFIALPRESNITVLIKLGPTLLAKPCYIV 60
Qy      61  SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGCPFGEVQLQPSLSLPT 120
Db      61  SKRHITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMGCPFGEVQLQPSLSLPT 120
Qy      121  LNRFTIWDVKAKHSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db      121  LNRFTIWDVKAKHSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Qy      121  LNRFTIWDVKAKHSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db      121  LNRFTIWDVKAKHSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Qy      181  GTVSRIKMOEGVKMALHLPMFHPNRVSGFSIANRSSIKRLCIIESVFEGSGSATLMSANY 240
Db      181  GTVSRIKMOEGVKMALHLPMFHPNRVSGFSIANRSSIKRLCIIESVFEGSGSATLMSANY 240
Qy      241  PEGFPEDELMTWQFVVPAPHLRASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
Db      241  PEGFPEDELMTWQFVVPAPHLRASVSFLNPNLSNCRKERVEYYIPGSTTNPVEVKLBK 300
Qy      301  QPGNMGNFNLSLQGCDDAOSPGLRLQFOVLVOHPONESNK 343
Db      301  QPGNMGNFNLSLQGCDDAOSPGLRLQFOVLVOHPONESSE 343

RESULT 13
US-10-176-918-162
; Sequence 162, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
```

APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C382
CURRENT APPLICATION NUMBER: US/10/176,918
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-918-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4,1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLIGAAARLPFGAFAFEIALPRESNITVLIKLGPTLLAKPCYIVI 60
DB 1 MAGLNCGVSIALLGVLLIGAAARLPFGAFAFEIALPRESNITVLIKLGPTLLAKPCYIVI 60
QY 61 SKRHITMLSIKSGRIVTFSCQSPENHFVIEIQKNIDCMGCPFGFVQLOPSTSLPT 120
DB 61 SKRHITMLSIKSGRIVTFSCQSPENHFVIEIQKNIDCMGCPFGFVQLOPSTSLPT 120
QY 121 LNRFTIMDVKAHKSIGLEQFSIPRLROIGGESCPDGVTHSIGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLEQFSIPRLROIGGESCPDGVTHSIGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKQEGVKALHPWFHPRNVSQFSIANRSSIKRLCTIESVFEGESATLMSANY 240
DB 181 GTVSRIKQEGVKALHPWFHPRNVSQFSIANRSSIKRLCTIESVFEGESATLMSANY 240
QY 241 PEGFPEDELMTWQFVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTNPVEYFKLEDK 300
DB 241 PEGFPEDELMTWQFVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTNPVEYFKLEDK 300
QY 301 QPGNMAGNFNLSLOGCDQDAQSPGILRLQFOVLVQHONESNK 343
DB 301 QPGNMAGNFNLSLOGCDQDAQSPGILRLQFOVLVQHONESNK 343

RESULT 14
US-10-176-921-162
Sequence 162, Application US/10176921
Publication No. US20030027276A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C288
CURRENT APPLICATION NUMBER: US/10/176,921
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-921-162

Query Match 40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4,1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLIGAAARLPFGAFAFEIALPRESNITVLIKLGPTLLAKPCYIVI 60
DB 1 MAGLNCGVSIALLGVLLIGAAARLPFGAFAFEIALPRESNITVLIKLGPTLLAKPCYIVI 60
QY 61 SKRHITMLSIKSGRIVTFSCQSPENHFVIEIQKNIDCMGCPFGFVQLOPSTSLPT 120
DB 61 SKRHITMLSIKSGRIVTFSCQSPENHFVIEIQKNIDCMGCPFGFVQLOPSTSLPT 120
QY 121 LNRFTIMDVKAHKSIGLEQFSIPRLROIGGESCPDGVTHSIGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLEQFSIPRLROIGGESCPDGVTHSIGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKQEGVKALHPWFHPRNVSQFSIANRSSIKRLCTIESVFEGESATLMSANY 240
DB 181 GTVSRIKQEGVKALHPWFHPRNVSQFSIANRSSIKRLCTIESVFEGESATLMSANY 240
QY 241 PEGFPEDELMTWQFVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTNPVEYFKLEDK 300
DB 241 PEGFPEDELMTWQFVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTNPVEYFKLEDK 300
QY 301 QPGNMAGNFNLSLOGCDQDAQSPGILRLQFOVLVQHONESNK 343
DB 301 QPGNMAGNFNLSLOGCDQDAQSPGILRLQFOVLVQHONESNK 343

RESULT 15
US-10-137-865-162
Sequence 162, Application US/10137865
Publication No. US20030032155A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C154

```

; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-865-162

```

```

Query Match      40.6%; Score 1779; DB 4; Length 343;
Best Local Similarity 99.4%; Pred. No. 4.1e-150;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MAGLNCGVSIALLGVLLIGAARLPARGAEAFETALPRESNITVLIKGPTTLAKPCYIV 60
Db      1 MAGLNCGVSIALLGVLLIGAARLPARGAEAFETALPRESNITVLIKGPTTLAKPCYIV 60
QY      61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIRIOKNIDCMGSPCPFGEVQLQPTSLIPT 120
Db      61 SKRHITMLSIKSGERIVFTFSCQSPENHFVIRIOKNIDCMGSPCPFGEVQLQPTSLIPT 120
QY      121 INRTFIWDVKAKHSIGLEIQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db      121 INRTFIWDVKAKHSIGLEIQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY      181 GTVSRIKMOEGVKMALHLPMFHPRVNVSFSIANRSSIKRLCTIESVFESEGSATLMSANY 240
Db      181 GTVSRIKMOEGVKMALHLPMFHPRVNVSFSIANRSSIKRLCTIESVFESEGSATLMSANY 240
QY      241 PEGPPEDELMTWQFVVPAILRASVSFLNENLNCERKEKREVEYYIPGSTTNPEVFKLBK 300
Db      241 PEGPPEDELMTWQFVVPAILRASVSFLNENLNCERKEKREVEYYIPGSTTNPEVFKLBK 300
QY      301 QPGNMAGNFNLSLQGCDDQDQSPGILRLQFQVLVOHPQNESNK 343
Db      301 QPGNMAGNFNLSLQGCDDQDQSPGILRLQFQVLVOHPQNESSE 343

```

Search completed: May 4, 2006, 22:31:46
 Job time : 86 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 22:32:01 ; Search time 17 Seconds

(without alignments) 2276.120 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4377

Sequence: 1 MAGNCGVSIALLGVLLGGA.....SSKTDIDPLXQEMERPAE 836

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 235405

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*
1: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB.pep1.*
2: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB.pep1.*
3: /SIDS5/ptodata/2/pubpaa/US07_NEW_PUB.pep1.*
4: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB.pep1.*
5: /SIDS5/ptodata/2/pubpaa/PCT_NEW_PUB.pep1.*
6: /SIDS5/ptodata/2/pubpaa/US09_NEW_PUB.pep1.*
7: /SIDS5/ptodata/2/pubpaa/US09_NEW_PUB.pep1.*
8: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
9: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
10: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep1.*
11: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep1.*
12: /SIDS5/ptodata/2/pubpaa/US60_NEW_PUB.pep1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4375	99.9	836	11	US-11-050-857-989
2	4374	99.9	836	11	US-11-050-857-988
3	4048	92.5	770	11	US-11-050-857-252
4	3625.5	82.8	709	11	US-11-229-769-132
5	3395	77.6	649	11	US-11-050-857-990
6	2597	59.3	495	11	US-11-050-857-1135
7	2168	49.5	487	11	US-11-050-857-253
8	2155	49.2	414	11	US-11-229-769-273
9	2155	49.2	443	11	US-11-229-769-271
10	1779	40.6	343	9	US-10-131-826A-162
11	1779	40.6	343	9	US-10-137-115B-162
12	1779	40.6	343	9	US-10-137-873A-162
13	1779	40.6	343	9	US-10-152-370-162
14	1779	40.6	343	11	US-11-056-857-987
15	1779	40.6	343	11	US-11-290-153-162
16	1210	27.6	231	11	US-11-050-857-254
17	1132	25.9	238	11	US-11-050-857-255
18	827	18.9	159	11	US-11-229-769-272
19	391	8.9	146	11	US-11-050-857-1136
20	138	3.2	4051	8	US-10-501-834-7
21	127	2.9	595	11	US-11-264-096-2258

22	118.5	2.7	4059	8	US-10-501-834-6	Sequence 6, Appl1
23	118	2.7	2766	9	US-10-877-346-62	Sequence 62, Appl1
24	117	2.7	3623	9	US-10-995-561-553	Sequence 593, Appl1
25	113	2.6	869	9	US-10-453-372-50	Sequence 50, Appl1
26	113	2.6	2612	9	US-10-453-372-38	Sequence 38, Appl1
27	113	2.6	2669	9	US-10-453-372-36	Sequence 36, Appl1
28	113	2.6	3095	11	US-11-235-732-4	Sequence 4, Appl1
29	113	2.6	3104	9	US-10-453-372-34	Sequence 34, Appl1
30	113	2.6	3104	9	US-10-453-372-62	Sequence 62, Appl1
31	113	2.6	3104	9	US-10-453-372-64	Sequence 64, Appl1
32	113	2.6	3130	9	US-10-453-372-42	Sequence 42, Appl1
33	113	2.6	3483	9	US-10-453-372-40	Sequence 40, Appl1
34	113	2.6	3546	9	US-10-453-372-32	Sequence 32, Appl1
35	112.5	2.6	1725	9	US-10-766-317-8	Sequence 8, Appl1
36	112.5	2.6	1725	9	US-10-784-004-457	Sequence 457, Appl1
37	112.5	2.6	1725	9	US-10-784-004-960	Sequence 960, Appl1
38	112	2.6	1045	11	US-11-096-568A-30402	Sequence 30402, A
39	108	2.5	699	11	US-11-150-887-43	Sequence 43, Appl1
40	106.5	2.4	1125	11	US-11-024-959-160	Sequence 360, Appl1
41	105	2.4	4590	8	US-10-505-928-569	Sequence 569, Appl1
42	104	2.4	766	11	US-11-189-301-21	Sequence 21, Appl1
43	104	2.4	2768	9	US-10-510-101-72	Sequence 72, Appl1
44	101.5	2.3	693	11	US-11-189-301-20	Sequence 20, Appl1
45	101.5	2.3	927	11	US-11-189-301-10	Sequence 10, Appl1

ALIGNMENTS

RESULT 1																																
US-11-050-857-989																																
; Sequence 989, Application US/11050857																																
; Publication No. US20060040278A1																																
; GENERAL INFORMATION:																																
; APPLICANT: Compugen Ltd																																
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS																																
; FILE REFERENCE: 1847.1005																																
; CURRENT APPLICATION NUMBER: US/11/050.857																																
; CURRENT FILING DATE: 2005-01-27																																
; NUMBER OF SEQ ID NOS: 1150																																
; SEQ ID NO 989																																
; LENGTH: 836																																
; TYPE: PRT																																
; ORGANISM: Homo sapiens																																
US-11-050-857-989																																
Query Match																																
Best Local Similarity 99.6%; Score 4375; DB 11; Length 836;																																
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;																																
QY	1	MAGNCGVSIALLGVLLGGA	RLPRGA	AFRIAL	PRESENT	IVL	IKL	GPTTL	LAKPCY	VI 60																						
DB	1	MAGNCGVSIALLGVLLGGA	RLPRGA	AFRIAL	PRESENT	IVL	IKL	GPTTL	LAKCY	VI 60																						
QY	61	SKRHITMISIGSERIVFTFSCQSPENH	FVIEIQ	NIDMS	GCP	FGV	LO	PTSL	LPT	120																						
DB	61	SKRHITMISIGSERIVFTFSCQSPENH	FVIEIQ	NIDMS	GCP	FGV	LO	PTSL	LPT	120																						
QY	121	LNRFITMVKHKSIGLEOF	SIPLR	IOIG	SGSC	PDV	TS	ISGR	IDA	TVR	IGT	PCSN 180																				
DB	121	LNRFITMVKHKSIGLEOF	SIPLR	IOIG	SGSC	PDV	TS	ISGR	IDA	TVR	IGT	PCSN 180																				
QY	181	GTGVRIRKQGVKALHL	PMFPR	NVSG	FS	IANK	SSIR	LC	II	SV	FG	SG	SAT	LMS	ANY 240																	
DB	181	GTGVRIRKQGVKALHL	PMFPR	NVSG	FS	IANK	SSIR	LC	II	SV	FG	SG	SAT	LMS	ANY 240																	
QY	241	PEGFPEDELMTQGV	PAHL	RA	SV	PLNF	INS	CER	KE	RE	V	Y	I	PG	ST	N	PE	V	K	LEDK 300												
DB	241	PEGFPEDELMTQGV	PAHL	RA	SV	PLNF	INS	CER	KE	RE	V	Y	I	PG	ST	N	PE	V	K	LEDK 300												
QY	301	ORGWAGNFN	SL	GG	CD	AD	AS	PG	IL	R	Q	PO	V	Y	OH	P	ON	EN	K	I	Y	Y	U	D	L	S	R	M	S	L	T	IE 360
DB	301	ORGWAGNFN	SL	GG	CD	AD	AS	PG	IL	R	Q	PO	V	Y	OH	P	ON	EN	K	I	Y	Y	U	D	L	S	R	M	S	L	T	IE 360

```
Db 301 QPGNAGNFNLSTGDCDDAOSPGLRLQFVLVQHPPONESNKIYVVDLSNBRASLTIE 360
Qy 361 PRPVQSKRPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDDLTRLMMNVEKTIISCTDH 420
Db 361 PRPVQSKRPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDDLTRLMMNVEKTIISCTDH 420
Qy 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLVLPKORLSLVLPACKLOQHTHEKCNISF 480
Db 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLVLPKORLSLVLPACKLOQHTHEKCNISF 480
Qy 481 SYLVASAIPODLYGSCFCPGGSIKOIQVKONISVTLRTFAPSFYQEASRQGLTVSFLPY 540
Db 481 SYLVASAIPODLYGSCFCPGGSIKOIQVKONISVTLRTFAPSFYQEASRQGLTVSFLPY 540
Qy 541 FKEBGFVTPTDYSKVYLRTPNMDRGPLSLTSVSWNISVPRDQVACLTFFKERSGVVCQ 600
Db 541 FKEBGFVTPTDYSKVYLRTPNMDRGPLSLTSVSWNISVPRDQVACLTFFKERSGVVCQ 600
Qy 601 TGRAFMIIOEORTRAEEIFSLDEDLVLPKPSFHHHSFWWNISNCSPTSGKQDLFLSVTLT 660
Db 601 TGRAFMIIOEORTRAEEIFSLDEDLVLPKPSFHHHSFWWNISNCSPTSGKQDLFLSVTLT 660
Qy 661 PRVVDLTVILAAVGGVLLLSALGLIICVKKKKKTKNGPAGVIYXNINTEMPROPK 720
Db 661 PRVVDLTVILAAVGGVLLLSALGLIICVKKKKKTKNGPAGVIYXNINTEMPROPK 720
Qy 721 KFQGRKNDSHVAVIEDTWYGHLLQDSSGSFLQPEVDYTRPPQGTMGVCPSPPTIC 780
Db 721 KFQGRKNDSHVAVIEDTWYGHLLQDSSGSFLQPEVDYTRPPQGTMGVCPSPPTIC 780
Qy 781 SRAPAKLATEBPPRSPSPSESEBPTTSHPNNGVSSKDTDIPLLXTOEPMERAE 836
Db 781 SRAPAKLATEBPPRSPSPSESEBPTTSHPNNGVSSKDTDIPLLXTOEPMERAE 836
```

RESULT 2

```
US-11-050-857-988
; Sequence 988, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 988
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-988
```

```
Query Match 99.9%; Score 4374; DB 11; Length 836;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAGNCGVSIALGVLIGAAARLPAGAFAEIALPRESNITVLIKIGPTLLAKPCYVI 60
Db 1 MAGNCGVSIALGVLIGAAARLPAGAFAEIALPRESNITVLIKIGPTLLAKPCYVI 60
Qy 61 SKRHITMTSISKGERIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGVQLOPSTSLPT 120
Db 61 SKRHITMTSISKGERIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGVQLOPSTSLPT 120
Qy 121 INRTPIWDYKAHKSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCGN 180
Db 121 INRTPIWDYKAHKSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCGN 180
Qy 181 GTVSRIKQOEGVKALHLPMFHPRNVSQFSIANRSSIKRLCIISVFEGESATLMSANY 240
Db 181 GTVSRIKQOEGVKALHLPMFHPRNVSQFSIANRSSIKRLCIISVFEGESATLMSANY 240
```

```
Qy 241 PEGPEDELTMTQFVVPVPAHLRASVFLNFNLSCNCKERBEVEYYIIPGSTNPEVEKLEBK 300
Db 241 PEGPEDELTMTQFVVPVPAHLRASVFLNFNLSCNCKERBEVEYYIIPGSTNPEVEKLEBK 300
Qy 301 QPGNAGNFNLSTGDCDDAOSPGLRLQFVLVQHPPONESNKIYVVDLSNBRASLTIE 360
Db 301 QPGNAGNFNLSTGDCDDAOSPGLRLQFVLVQHPPONESNKIYVVDLSNBRASLTIE 360
Qy 361 PRPVQSKRPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDDLTRLMMNVEKTIISCTDH 420
Db 361 PRPVQSKRPVPGCFVCLIESRTCSNLTLTSGSKHKISFLCDDLTRLMMNVEKTIISCTDH 420
Qy 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLVLPKORLSLVLPACKLOQHTHEKCNISF 480
Db 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLVLPKORLSLVLPACKLOQHTHEKCNISF 480
Qy 481 SYLVASAIPODLYGSCFCPGGSIKOIQVKONISVTLRTFAPSFYQEASRQGLTVSFLPY 540
Db 481 SYLVASAIPODLYGSCFCPGGSIKOIQVKONISVTLRTFAPSFYQEASRQGLTVSFLPY 540
Qy 541 FKEBGFVTPTDYSKVYLRTPNMDRGPLSLTSVSWNISVPRDQVACLTFFKERSGVVCQ 600
Db 541 FKEBGFVTPTDYSKVYLRTPNMDRGPLSLTSVSWNISVPRDQVACLTFFKERSGVVCQ 600
Qy 601 TGRAFMIIOEORTRAEEIFSLDEDLVLPKPSFHHHSFWWNISNCSPTSGKQDLFLSVTLT 660
Db 601 TGRAFMIIOEORTRAEEIFSLDEDLVLPKPSFHHHSFWWNISNCSPTSGKQDLFLSVTLT 660
Qy 661 PRVVDLTVILAAVGGVLLLSALGLIICVKKKKKTKNGPAGVIYXNINTEMPROPK 720
Db 661 PRVVDLTVILAAVGGVLLLSALGLIICVKKKKKTKNGPAGVIYXNINTEMPROPK 720
Qy 721 KFQGRKNDSHVAVIEDTWYGHLLQDSSGSFLQPEVDYTRPPQGTMGVCPSPPTIC 780
Db 721 KFQGRKNDSHVAVIEDTWYGHLLQDSSGSFLQPEVDYTRPPQGTMGVCPSPPTIC 780
Qy 781 SRAPAKLATEBPPRSPSPSESEBPTTSHPNNGVSSKDTDIPLLXTOEPMERAE 836
Db 781 SRAPAKLATEBPPRSPSPSESEBPTTSHPNNGVSSKDTDIPLLXTOEPMERAE 836
```

RESULT 3

```
US-11-050-857-252
; Sequence 252, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 252
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-252
```

```
Query Match 92.5%; Score 4048; DB 11; Length 770;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 767; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 67 MUISKGERIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGVQLOPSTSLPTLNRTFI 126
Db 1 MUISKGERIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGVQLOPSTSLPTLNRTFI 126
Qy 127 MDVKAHKSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCNSGVSR 186
Db 61 MDVKAHKSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCNSGVSR 120
Qy 187 KMQOEGVKALHLPMFHPRNVSQFSIANRSSIKRLCIISVFEGESATLMSANYEGPPE 246
```

```

Db      121 KMGGVAKMALHPMFHPNNGVGFSTANSSIKRLCTIISVEGEGSATLMSANYPEGRPE 180
Qy      247 DELMTQFVVAHLRASVSLNPNLSNCRKEERYEYI PGSTTNPEVFKLEDKQPGMA 306
Db      181 DELMTQFVVAHLRASVSLNPNLSNCRKEERYEYI PGSTTNPEVFKLEDKQPGMA 240
Qy      307 GNFMLSLQGCCDDQASPGILRLQOVLVQHONSNKLYVVDLSNERAMSLTIEPRPVKQ 366
Db      241 GNFMLSLQGCCDDQASPGILRLQOVLVQHONSNKLYVVDLSNERAMSLTIEPRPVKQ 300
Qy      367 SRKRVPCGFVCLERTCSNLTITSGSGHKISFLCDDTLRLMMVVEKTIISCTDHRVCOR 426
Db      301 SRKRVPCGFVCLERTCSNLTITSGSGHKISFLCDDTLRLMMVVEKTIISCTDHRVCOR 360
Qy      427 SYSIQVPSDILHLPEVLFHDFSMKLLVPKDRLSLVLPKQKIQOHTHEKPCNTSFSYLVAS 486
Db      361 SYSIQVPSDILHLPEVLFHDFSMKLLVPKDRLSLVLPKQKIQOHTHEKPCNTSFSYLVAS 420
Qy      487 AIPBODLYFGSFCGSGSIKOIOVKONISVTLRTAPASXQASRQGLTVSPITPYKKEGV 546
Db      421 AIPBODLYFGSFCGSGSIKOIOVKONISVTLRTAPASXQASRQGLTVSPITPYKKEGV 480
Qy      547 FTVTPDTRSKYLLRTPMWDRGLPSLTSVSMNISVPRDVACLTFPKERSGVVCOGRAPFM 606
Db      481 FTVTPDTRSKYLLRTPMWDRGLPSLTSVSMNISVPRDVACLTFPKERSGVVCOGRAPFM 540
Qy      607 IIOGRTAEIEIFSLDEVDLPKPSFHHHSFWVNISNCSPTSGKQDLDFSVTLPRTVDL 666
Db      541 IIOGRTAEIEIFSLDEVDLPKPSFHHHSFWVNISNCSPTSGKQDLDFSVTLPRTVDL 600
Qy      667 TVIIIAVGGVLLLSALGLICCVKKKKKTKNGPANGIYNXNINTEMPQPKPKQGR 726
Db      601 TVIIIAVGGVLLLSALGLICCVKKKKKTKNGPANGIYNXNINTEMPQPKPKQGR 660
Qy      727 KDNDSHYAVVEDTMVYGHLLQDSSGFLQPEVDTPRFGOTMGVCPSPPTISRAPTA 786
Db      661 KDNDSHYAVVEDTMVYGHLLQDSSGFLQPEVDTPRFGOTMGVCPSPPTISRAPTA 720
Qy      787 KLAIEPPRPSPESESEPYTFSHPNNGDVSSKDTDIPLLXTOEPMEDAE 836
Db      721 KLAIEPPRPSPESESEPYTFSHPNNGDVSSKDTDIPLLXTOEPMEDAE 770

```

```

RESULT 4
US-11-229-769-132
; Sequence 132, Application US/11229769
; Publication No. US20060079670A1
; GENERAL INFORMATION:
; APPLICANT: Komatsugai et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1D1C1
; CURRENT APPLICATION NUMBER: US/11/229, 769
; CURRENT FILING DATE: 2005-09-20
; PRIOR APPLICATION NUMBER: 10/233,453
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 132
; LENGTH: 709

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (189)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (275)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (414)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (438)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (641)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (643)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (696)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (697)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-11-229-769-132

Query Match      82.8%; Score 3625.5; DB 11; Length 709;
Best Local Similarity 98.3%; Pred. No. 2,1e-308;
Matches 698; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

Qy      1 MAGNLCGVSTALGLVLLGGAARLPRGAEPFIALPRESNTITVLKIGPTLLAKPCYIVI 60
Db      1 MAGNLCGVSTALGLVLLGGAARLPRGAEPFIALPRESNTITVLKIGPTLLAKPCYIVI 60
Qy      61 SKRHITMLSIKSGRIIVFTSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSLSLPT 120
Db      61 SKRHITMLSIKSGRIIVFTSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSLSLPT 120
Qy      121 LNRFTFIDVYAKHSIGLEQFSIRLRLQIIGBESCPGVTHSIGRIDATVVRIGTFCSN 180
Db      121 LNRFTFIDVYAKHSIGLEQFSIRLRLQIIGBESCPGVTHSIGRIDATVVRIGTFCSN 180
Qy      181 GTVSRIMOGGVKMALHPMFHPNNGVGFSTANSSIKRLCTIISVEGEGSATLMSANY 240
Db      181 GTVSRIMOGGVKMALHPMFHPNNGVGFSTANSSIKRLCTIISVEGEGSATLMSANY 240
Qy      241 PEGPPEDELMTQFVVAHLRASVSLNPNLSNCRKEERYEYI PGSTTNPEVFKLEDK 300
Db      241 PEGPPEDELMTQFVVAHLRASVSLNPNLSNCRKEERYEYI PGSTTNPEVFKLEDK 300
Qy      301 QPGNMAGNPNLSLOGCDDQASPGILRLQOVLVQHONSNKLYVVDLSNERAMSLTIE 360
Db      301 QPGNMAGNPNLSLOGCDDQASPGILRLQOVLVQHONSNKLYVVDLSNERAMSLTIE 360
Qy      361 PRPVKQSKRVPCGFVCLERTCSNLTITSGSGHKISFLCDDTLRLMMVVEKTIISCTDH 420
Db      361 PRPVKQSKRVPCGFVCLERTCSNLTITSGSGHKISFLCDDTLRLMMVVEKTIISCTDH 420
Qy      421 RYCQKRSYSIQVPSDILHLPEVLFHDFSMKLLVPKDRLSLVLPKQKIQOHTHEKPCNTSF 480
Db      421 RYCQKRSYSIQVPSDILHLPEVLFHDFSMKLLVPKDRLSLVLPKQKIQOHTHEKPCNTSF 480
Qy      481 SYLVASAIPOBODLYFGSFCGSGSIKOIOVKONISVTLRTAPASXQASRQGLTVSPITPY 540

```

Db 481 SYLVASALPSODLYGSCFCPGSIIQIYKONISVTLRTAFSPFOEASRQGLTFSIFPY 540
Qy 541 FKEBSEVFTVPTPTKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFEKERSGVVQ 600
Db 541 FKEBSEVFTVPTPTKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFEKERSGVVQ 600
Qy 601 TGRAMIIQOEKTRABEIFSLEDEVLPRKSFHHHSFWNISCSPTSGKQDLRSVTLT 660
Db 601 TGRAMIIQOEKTRABEIFSLEDEVLPRKSFHHHSFWNISCSPTSGKQDLRSVTLT 660
Qy 661 PRVVDLTVILAAVGGVLLSALGLIICVKKKKKKTKNKGPAVGIIYNN 710
Db 661 PRVVDLTVILAAVGGVLLSALGLIICVKKKKKKTKNKGPAVGIIYNN 709
RESULT 5
US-11-050-857-990
; Sequence 990, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Comugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; TITLE OF INVENTION: THEREOF FOR DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 990
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-990

Query Match 77.6%; Score 3395; DB 11; Length 649;
Best Local Similarity 99.2%; Pred. No. 2,6e-288;
Matches 644; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 188 MOEGKMLHLPMFPRNVSGFSIANRSIKRLCTIESFEGEGATLMSAYPGEPED 247
Db 1 MOEGKMLHLPMFPRNVSGFSIANRSIKRLCTIESFEGEGATLMSAYPGEPED 60
Qy 248 ELMTQFVPAHLRASVSFLNPNLSNCRKEERVEYITPGSTINPEVFLBDKQGNMAG 307
Db 61 ELMTQFVPAHLRASVSFLNPNLSNCRKEERVEYITPGSTINPEVFLBDKQGNMAG 120
Qy 308 NFNISLQCCDDAQSFGILRLQFOYLVQHPONESNKIYVVDISNERAMSLTEPRPVQS 367
Db 121 NFNISLQCCDDAQSFGILRLQFOYLVQHPONESNKIYVVDISNERAMSLTEPRPVQS 180
Qy 368 RKFVGCFCVLESRTCCSNLTLTSGSKHISFLCDDLRLMNNVEKTIISCTDHRVCORKS 427
Db 181 RKFVGCFCVLESRTCCSNLTLTSGSKHISFLCDDLRLMNNVEKTIISCTDHRVCORKS 240
Qy 428 YSLQVPSDILHLPELHDFSMKLLVPRKRLSLVLPAPQKLOQHTHEKPCNTSFSYLVA 487
Db 241 YSLQVPSDILHLPELHDFSMKLLVPRKRLSLVLPAPQKLOQHTHEKPCNTSFSYLVA 300
Qy 488 IPSQDLTFGSGFCPGSIIQIYKONISVTLRTAFSPFOEASRQGLTFSIFYEKEGV 547
Db 301 IPSQDLTFGSGFCPGSIIQIYKONISVTLRTAFSPFOEASRQGLTFSIFYEKEGV 360
Qy 548 TTPPTKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFEKERSGVVQOTGRAFMI 607
Db 361 TTPPTKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFEKERSGVVQOTGRAFMI 420
Qy 608 IOEQRTRABEIFSLEDEVLPRKSFHHHSFWNISCSPTSGKQDLRSVTLTPTVDT 667
Db 421 IOEQRTRABEIFSLEDEVLPRKSFHHHSFWNISCSPTSGKQDLRSVTLTPTVDT 480
Qy 668 VILAAVGGVLLSALGLIICVKKKKKKTKNKGPAVGIIYNNINTEMPRQPKFOKGRK 727
Db 481 VILAAVGGVLLSALGLIICVKKKKKKTKNKGPAVGIIYNNINTEMPRQPKFOKGRK 540

Qy 728 DNDSHVAVIEDTWYGHLLQDSSGSFLQPEVDTRYRPOGTMGVCPSPPTICSRAPTAK 787
Db 541 DNDSHVAVIEDTWYGHLLQDSSGSFLQPEVDTRYRPOGTMGVCPSPPTICSRAPTAK 600
Qy 788 LATEPPRSPPESESEPTFSHPNNGVSSKQDIDPLLXTOEPMEPAE 836
Db 601 LATEPPRSPPESESEPTFSHPNNGVSSKQDIDPLLXTOEPMEPAE 649
RESULT 6
US-11-050-857-1135
; Sequence 1135, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Comugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; TITLE OF INVENTION: THEREOF FOR DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 1135
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-11-050-857-1135

Query Match 59.3%; Score 2597; DB 11; Length 495;
Best Local Similarity 99.4%; Pred. No. 1e-218;
Matches 492; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 342 NKIYVVDISNERAMSLTEPRPVQKSRKFVGCFCVLESRTCCSNLTLTSGSKHISFLC 401
Db 1 NKIYVVDISNERAMSLTEPRPVQKSRKFVGCFCVLESRTCCSNLTLTSGSKHISFLC 60
Qy 402 DDLRLMNNVEKTIISCTDHRVCORKSYSLOVPSDILHLPELHDFSMKLLVPRKRLSLV 461
Db 61 DDLRLMNNVEKTIISCTDHRVCORKSYSLOVPSDILHLPELHDFSMKLLVPRKRLSLV 120
Qy 462 VPAQKLOQHTHEKPCNTSFSYLVAALPSQDLTFGSGFCPGSIIQIYKONISVTLRTFA 521
Db 121 VPAQKLOQHTHEKPCNTSFSYLVAALPSQDLTFGSGFCPGSIIQIYKONISVTLRTFA 180
Qy 522 PSFYQEASRQGLTFSIFYEKEGVFTVPTPTKSKVYLRTPNMDRGLPSLTSVSNISVP 581
Db 181 PSFYQEASRQGLTFSIFYEKEGVFTVPTPTKSKVYLRTPNMDRGLPSLTSVSNISVP 240
Qy 582 RDQVACLTFEKERSGVVQOTGRAFMIIOEQRTRABEIFSLEDEVLPRKSFHHHSFWNIS 641
Db 241 RDQVACLTFEKERSGVVQOTGRAFMIIOEQRTRABEIFSLEDEVLPRKSFHHHSFWNIS 300
Qy 642 NCSPTSGKQDLRSVTLTPTVDTVILAAVGGVLLSALGLIICVKKKKKKTKNG 701
Db 301 NCSPTSGKQDLRSVTLTPTVDTVILAAVGGVLLSALGLIICVKKKKKKTKNG 360
Qy 702 PAVGIIYNNINTEMPRQPKFOKGRKNDSHVAVIETTWYGHLLQDSSGSFLOPEVDT 761
Db 361 PAVGIIYNNINTEMPRQPKFOKGRKNDSHVAVIETTWYGHLLQDSSGSFLOPEVDT 420
Qy 762 YRPOGTMGVCPSPPTICSRAPTAKLATEPPRSPPESESEPTFSHPNNGVSSKQD 821
Db 421 YRPOGTMGVCPSPPTICSRAPTAKLATEPPRSPPESESEPTFSHPNNGVSSKQD 480
Qy 822 DPLLXTOEPMEPAE 836
Db 481 DPLLXTOEPMEPAE 495

RESULT 7
US-11-050-857-253

```
; Sequence 253, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 253
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-253

Query Match      49.5%; Score 2168; DB 11; Length 487;
Best Local Similarity 100.0%; Pred. No. 3,2e-181;
Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAGNCGVSIALDGVLLGAARLPRGAFAFIAPRESNITVLIKLGTPTLAKPCYIV 60
Db      1  MAGNCGVSIALDGVLLGAARLPRGAFAFIAPRESNITVLIKLGTPTLAKPCYIV 60
Qy      61  SKRHITMSTKSGRIYFTFSCQSPENHFVIEIKNIDCMGSPCPFGVQLOPSTSLPT 120
Db      61  SKRHITMSTKSGRIYFTFSCQSPENHFVIEIKNIDCMGSPCPFGVQLOPSTSLPT 120
Qy      121  LNRFTIMDVKAHKSIGLEQFSIPRLQIGGESCPCDVTHSISGRIDATVVRIGTFCSN 180
Db      121  LNRFTIMDVKAHKSIGLEQFSIPRLQIGGESCPCDVTHSISGRIDATVVRIGTFCSN 180
Qy      181  GTVSRIRKQEGVKALHLPWFHPNVSQFSIANRSSIRLCTIISVFEGESATLMSANY 240
Db      181  GTVSRIRKQEGVKALHLPWFHPNVSQFSIANRSSIRLCTIISVFEGESATLMSANY 240
Qy      241  PEGPEDELMTWQVVPVPAHLASVFLNPNLSNCRKEEREYIIPGSTNPEVFKEDK 300
Db      241  PEGPEDELMTWQVVPVPAHLASVFLNPNLSNCRKEEREYIIPGSTNPEVFKEDK 300
Qy      301  QPGMAGNFNLSLOGCDQDQSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE 360
Db      301  QPGMAGNFNLSLOGCDQDQSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE 360
Qy      361  PRPVKSRKFPVPGCFVCLIESRTCSNLTITSGSKHKISFLCDDLTRLMMNVEKTIS 416
Db      361  PRPVKSRKFPVPGCFVCLIESRTCSNLTITSGSKHKISFLCDDLTRLMMNVEKTIS 416

RESULT 8
US-11-229-769-273
; Sequence 273, Application US/11229769
; Publication No. US20060079670A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PIDICI
; CURRENT APPLICATION NUMBER: US/11/229,769
; PRIOR FILING DATE: 2005-09-20
; PRIOR APPLICATION NUMBER: 10/233,453
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
```

```
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 273
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-229-769-273

Query Match      49.2%; Score 2155; DB 11; Length 414;
Best Local Similarity 100.0%; Pred. No. 3,4e-180;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAGNCGVSIALDGVLLGAARLPRGAFAFIAPRESNITVLIKLGTPTLAKPCYIV 60
Db      1  MAGNCGVSIALDGVLLGAARLPRGAFAFIAPRESNITVLIKLGTPTLAKPCYIV 60
Qy      61  SKRHITMSTKSGRIYFTFSCQSPENHFVIEIKNIDCMGSPCPFGVQLOPSTSLPT 120
Db      61  SKRHITMSTKSGRIYFTFSCQSPENHFVIEIKNIDCMGSPCPFGVQLOPSTSLPT 120
Qy      121  LNRFTIMDVKAHKSIGLEQFSIPRLQIGGESCPCDVTHSISGRIDATVVRIGTFCSN 180
Db      121  LNRFTIMDVKAHKSIGLEQFSIPRLQIGGESCPCDVTHSISGRIDATVVRIGTFCSN 180
Qy      181  GTVSRIRKQEGVKALHLPWFHPNVSQFSIANRSSIRLCTIISVFEGESATLMSANY 240
Db      181  GTVSRIRKQEGVKALHLPWFHPNVSQFSIANRSSIRLCTIISVFEGESATLMSANY 240
Qy      241  PEGPEDELMTWQVVPVPAHLASVFLNPNLSNCRKEEREYIIPGSTNPEVFKEDK 300
Db      241  PEGPEDELMTWQVVPVPAHLASVFLNPNLSNCRKEEREYIIPGSTNPEVFKEDK 300
Qy      301  QPGMAGNFNLSLOGCDQDQSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE 360
Db      301  QPGMAGNFNLSLOGCDQDQSPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLTIE 360
Qy      361  PRPVKSRKFPVPGCFVCLIESRTCSNLTITSGSKHKISFLCDDLTRLMMNVEK 413
Db      361  PRPVKSRKFPVPGCFVCLIESRTCSNLTITSGSKHKISFLCDDLTRLMMNVEK 413

RESULT 9
US-11-229-769-271
; Sequence 271, Application US/11229769
; Publication No. US20060079670A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PIDICI
; CURRENT APPLICATION NUMBER: US/11/229,769
; PRIOR FILING DATE: 2005-09-20
; PRIOR APPLICATION NUMBER: 10/233,453
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 271
; LENGTH: 443
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-229-769-271

Query Match          49.2%; Score 2155; DB 11; Length 443;
Best Local Similarity 100.0%; Pred. No. 3,8e-180;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPFGAEAFETALPRESNITVLIKGTPTLLAKPCYVI 60
DB 30 MAGNCGVSIALGLVLLGAARLPFGAEAFETALPRESNITVLIKGTPTLLAKPCYVI 89
QY 61 SKRHITMLSIKSGERIVTFPSCQSPENHFVIEIQKNIDCMGCPGFGVQLOPSTSLPT 120
DB 90 SKRHITMLSIKSGERIVTFPSCQSPENHFVIEIQKNIDCMGCPGFGVQLOPSTSLPT 149
QY 121 LNRFTIMVKAHKSIGLELOFSIPRLROI GPESCPDGVTSHISGRIDATVIRIGTFCSN 180
DB 150 LNRFTIMVKAHKSIGLELOFSIPRLROI GPESCPDGVTSHISGRIDATVIRIGTFCSN 209
QY 161 GTVSRIKMOEGVKMALHPFHPHPRNVSGFSIANRSSIKRLCTIESVFEGSGATLMSANY 240
DB 210 GTVSRIKMOEGVKMALHPFHPHPRNVSGFSIANRSSIKRLCTIESVFEGSGATLMSANY 269
QY 241 PEGFPEDELMTQFVVPAPHLRASVSFLNPNLSNCRKERVRYIIPGSTTNEVFRLBDK 300
DB 270 PEGFPEDELMTQFVVPAPHLRASVSFLNPNLSNCRKERVRYIIPGSTTNEVFRLBDK 329
QY 301 QPGNAGNFNLSLQGCDDQDQSPGILRLQFOVLVOHPONESNKIYVVDLSNERAMSLTTE 360
DB 330 QPGNAGNFNLSLQGCDDQDQSPGILRLQFOVLVOHPONESNKIYVVDLSNERAMSLTTE 389
QY 361 PRPVQSKRFVPGFCVCLSESTCSNLTITSGSKHKISFLCDLTRLMANVEX 413
DB 390 PRPVQSKRFVPGFCVCLSESTCSNLTITSGSKHKISFLCDLTRLMANVEX 442

RESULT 10
US-10-131-826A-162
; Sequence 162, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P333ORIC128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117

```

```

; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-162

Query Match          40.6%; Score 1779; DB 9; Length 343;
Best Local Similarity 99.4%; Pred. No. 1.9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGLVLLGAARLPFGAEAFETALPRESNITVLIKGTPTLLAKPCYVI 60
DB 1 MAGNCGVSIALGLVLLGAARLPFGAEAFETALPRESNITVLIKGTPTLLAKPCYVI 60
QY 61 SKRHITMLSIKSGERIVTFPSCQSPENHFVIEIQKNIDCMGCPGFGVQLOPSTSLPT 120
DB 61 SKRHITMLSIKSGERIVTFPSCQSPENHFVIEIQKNIDCMGCPGFGVQLOPSTSLPT 120
QY 121 LNRFTIMVKAHKSIGLELOFSIPRLROI GPESCPDGVTSHISGRIDATVIRIGTFCSN 180
DB 121 LNRFTIMVKAHKSIGLELOFSIPRLROI GPESCPDGVTSHISGRIDATVIRIGTFCSN 180
QY 161 GTVSRIKMOEGVKMALHPFHPHPRNVSGFSIANRSSIKRLCTIESVFEGSGATLMSANY 240
DB 181 GTVSRIKMOEGVKMALHPFHPHPRNVSGFSIANRSSIKRLCTIESVFEGSGATLMSANY 240
QY 241 PEGFPEDELMTQFVVPAPHLRASVSFLNPNLSNCRKERVRYIIPGSTTNEVFRLBDK 300
DB 241 PEGFPEDELMTQFVVPAPHLRASVSFLNPNLSNCRKERVRYIIPGSTTNEVFRLBDK 300
QY 301 QPGNAGNFNLSLQGCDDQDQSPGILRLQFOVLVOHPONESNK 343
DB 301 QPGNAGNFNLSLQGCDDQDQSPGILRLQFOVLVOHPONESNK 343

RESULT 11
US-10-973-115B-162
; Sequence 162, Application US/10973115B
; Publication No. US20060040351A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODIN
; FILE REFERENCE: 39870-333ORIC300C1

```

```

; CURRENT APPLICATION NUMBER: US/10/973,115B
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 10/145,747
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/028,072
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/581,742
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: PCT/US00/05746
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/135,736
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 60/123,090
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-973-115B-162

Query Match          40.6%; Score 1779; DB 9; Length 343;
Best Local Similarity 99.4%; Pred. No. 1,9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGVLLGAARLPGAEAFIALPRESNITVLIKGTPTLLAKPCYIV 60
DB 1 MAGNCGVSIALGVLLGAARLPGAEAFIALPRESNITVLIKGTPTLLAKPCYIV 60
QY 61 SKRHITMSTISGGERIVTFSCQSPENHFVIEIQKIDMSGPCPFGEVOLQPSLSLPT 120
DB 61 SKRHITMSTISGGERIVTFSCQSPENHFVIEIQKIDMSGPCPFGEVOLQPSLSLPT 120
QY 121 LNRFTFIMVKAHKSIGLELQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTFIMVKAHKSIGLELQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCIIESVFEGESATLMSANY 240
DB 181 GTVSRIKMQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCIIESVFEGESATLMSANY 240
QY 241 PEGPPEDELMTWQFVVPALHSAVSFLNPNLSNCRKEKEVEYIIPGTTNPEVFKLBDK 300
DB 241 PEGPPEDELMTWQFVVPALHSAVSFLNPNLSNCRKEKEVEYIIPGTTNPEVFKLBDK 300
QY 301 QPGMAGNFNLSTLQCCDDAOSPGILRLQFOVLVOHPONESNK 343
DB 301 QPGMAGNFNLSTLQCCDDAOSPGILRLQFOVLVOHPONESNK 343

RESULT 12
US-10-137-873A-162
; Sequence 162, Application US/10137873A
; Publication No. US20060084138a1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Deenoysers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William

```

```

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C149
; CURRENT APPLICATION NUMBER: US/10/137,873A
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: 60/045911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-137-873A-162

Query Match          40.6%; Score 1779; DB 9; Length 343;
Best Local Similarity 99.4%; Pred. No. 1,9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALGVLLGAARLPGAEAFIALPRESNITVLIKGTPTLLAKPCYIV 60
DB 1 MAGNCGVSIALGVLLGAARLPGAEAFIALPRESNITVLIKGTPTLLAKPCYIV 60
QY 61 SKRHITMSTISGGERIVTFSCQSPENHFVIEIQKIDMSGPCPFGEVOLQPSLSLPT 120
DB 61 SKRHITMSTISGGERIVTFSCQSPENHFVIEIQKIDMSGPCPFGEVOLQPSLSLPT 120
QY 121 LNRFTFIMVKAHKSIGLELQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTFIMVKAHKSIGLELQFSIPRLROIIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCIIESVFEGESATLMSANY 240
DB 181 GTVSRIKMQEGVKALHLPMFHPRNVSQFSIANRSSIKRLCIIESVFEGESATLMSANY 240
QY 241 PEGPPEDELMTWQFVVPALHSAVSFLNPNLSNCRKEKEVEYIIPGTTNPEVFKLBDK 300
DB 241 PEGPPEDELMTWQFVVPALHSAVSFLNPNLSNCRKEKEVEYIIPGTTNPEVFKLBDK 300
QY 301 QPGMAGNFNLSTLQCCDDAOSPGILRLQFOVLVOHPONESNK 343
DB 301 QPGMAGNFNLSTLQCCDDAOSPGILRLQFOVLVOHPONESNK 343

RESULT 13
US-10-152-370-162
; Sequence 162, Application US/10152370
; Publication No. US20060084139a1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Deenoysers, Luc
; APPLICANT: Filvaroff, Ellen

```

```

; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C407
; CURRENT APPLICATION NUMBER: US/10/152,370
; CURRENT FILING DATE: 2002-05-21
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-152-370-162

Query Match          40.6%; Score 1779; DB 9; Length 343;
Best Local Similarity 99.4%; Pred. No. 1.9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALLGVLILGAARLPGAEAFETALPRESNITVLIKGPPTLAKCYIYI 60
DB 1 MAGNCGVSIALLGVLILGAARLPGAEAFETALPRESNITVLIKGPPTLAKCYIYI 60
QY 61 SKRHITMTSISGERIVFTFSCQSPENHFVIEIQKIDCMGSPCFGEVQLQPSLSLPT 120
DB 61 SKRHITMTSISGERIVFTFSCQSPENHFVIEIQKIDCMGSPCFGEVQLQPSLSLPT 120
QY 121 LNRFTIMDVKAHKSIGLELQFSIPRLRQIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLELQFSIPRLRQIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIKMQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKERVEYYIPGSTTNPDEVFKLEDK 300
DB 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKERVEYYIPGSTTNPDEVFKLEDK 300
QY 301 QPGNMGNFNLSLQGCDDQDQSPGILRLQFVLVQHPQNESNK 343
DB 301 QPGNMGNFNLSLQGCDDQDQSPGILRLQFVLVQHPQNESSE 343

RESULT 14
US-11-050-857-987
; Sequence 987, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 987
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-857-987

Query Match          40.6%; Score 1779; DB 11; Length 343;
```

```

; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C321
; CURRENT APPLICATION NUMBER: US/11/290,153
; CURRENT FILING DATE: 2005-11-30
; PRIOR APPLICATION NUMBER: US/10/146,728
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18

US-11-290-153-162
; Sequence 162, Application US/11290153
; Publication No. US20060073568A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Betesini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Flvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C321
; CURRENT APPLICATION NUMBER: US/11/290,153
; CURRENT FILING DATE: 2005-11-30
; PRIOR APPLICATION NUMBER: US/10/146,728
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18

QY 1 MAGNCGVSIALLGVLILGAARLPGAEAFETALPRESNITVLIKGPPTLAKCYIYI 60
DB 1 MAGNCGVSIALLGVLILGAARLPGAEAFETALPRESNITVLIKGPPTLAKCYIYI 60
QY 61 SKRHITMTSISGERIVFTFSCQSPENHFVIEIQKIDCMGSPCFGEVQLQPSLSLPT 120
DB 61 SKRHITMTSISGERIVFTFSCQSPENHFVIEIQKIDCMGSPCFGEVQLQPSLSLPT 120
QY 121 LNRFTIMDVKAHKSIGLELQFSIPRLRQIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDVKAHKSIGLELQFSIPRLRQIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIKMQEGVKALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKERVEYYIPGSTTNPDEVFKLEDK 300
DB 241 PEGFPEDELMTWQFVVPALRASVSFLNPNLSNCRKERVEYYIPGSTTNPDEVFKLEDK 300
QY 301 QPGNMGNFNLSLQGCDDQDQSPGILRLQFVLVQHPQNESNK 343
DB 301 QPGNMGNFNLSLQGCDDQDQSPGILRLQFVLVQHPQNESSE 343

RESULT 15
US-11-290-153-162
; Sequence 162, Application US/11290153
; Publication No. US20060073568A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Betesini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Flvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C321
; CURRENT APPLICATION NUMBER: US/11/290,153
; CURRENT FILING DATE: 2005-11-30
; PRIOR APPLICATION NUMBER: US/10/146,728
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18

Best Local Similarity 99.4%; Pred. No. 1.9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```


;; PRIOR APPLICATION NUMBER: 60/059352
;; PRIOR FILING DATE: 1997-09-19
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 162
;; LENGTH: 343
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-11-290-153-162

Query Match 40.6%; Score 1779; DB 11; Length 343;
Best Local Similarity 99.4%; Pred. No. 1.9e-147;
Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLIGAARLPRGAFAFEIALPRESNITVLIKLGPTLLAKPCYIVI 60
Db 1 MAGLNCGVSIALLGVLLIGAARLPRGAFAFEIALPRESNITVLIKLGPTLLAKPCYIVI 60
QY 61 SKRHITMLSIKSGRIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGEVQLPSTSLPT 120
Db 61 SKRHITMLSIKSGRIVFTFSCQSPENHFVIEIQKNIDMSGPCPFGEVQLPSTSLPT 120
QY 121 LNRFTIMDVKAHKSIGLELOFSIPRLQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTIMDVKAHKSIGLELOFSIPRLQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPMFHPRVNSGFSIANRSSIKRLCIIESVFEGEGSATLMSANY 240
Db 181 GTVSRIRKQEGVKALHLPMFHPRVNSGFSIANRSSIKRLCIIESVFEGEGSATLMSANY 240
QY 241 PEGFPEDELMTWQFVVPALRLASVSFLNPNLSNCRKEERVEYYIPGSTNPEVFKLEDK 300
Db 241 PEGFPEDELMTWQFVVPALRLASVSFLNPNLSNCRKEERVEYYIPGSTNPEVFKLEDK 300
QY 301 QPGNMAGNFNLSLQSCDDAOSPGIIRLQFQVLVQHPPONESNK 343
Db 301 QPGNMAGNFNLSLQSCDDAOSPGIIRLQFQVLVQHPPONESSE 343

Search completed: May 4, 2006, 22:32:43
Job time : 19 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model
Run on: May 4, 2006, 22:15:20 ; Search time 45 Seconds
(without alignments)
1787.494 Million cell updates/sec

Title: US-10-781-564-1
Perfect score: 4377
Sequence: 1 MAGINCGVSTALGVLLGA.....SSKOTDPLXTOEPMBAE 836

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	137.5	3.1	3623	2	T08618
2	121.5	2.8	986	2	T33135
3	118.5	2.7	2476	2	T34022
4	117.5	2.7	1464	2	S58984
5	117	2.7	2499	1	A30788
6	117	2.7	3623	2	T09456
7	113.5	2.6	1263	2	AH2011
8	113	2.6	1694	2	S50065
9	112.5	2.6	440	2	A39613
10	112	2.6	582	1	VCVDAR
11	112	2.6	1228	2	S46754
12	111.5	2.5	942	1	U01674
13	111	2.5	603	2	T24315
14	111	2.5	1185	2	T19212
15	111	2.5	3329	2	T42205
16	111	2.5	3329	2	T30904
17	111	2.5	4919	2	T31105
18	110.5	2.5	692	1	S59833
19	110.5	2.5	1027	2	T19173
20	110.5	2.5	1828	2	A40115
21	109.5	2.5	737	2	T31349
22	109	2.5	1224	2	T40765
23	108.5	2.5	5762	2	A41819
24	108	2.5	699	1	S47663
25	107.5	2.5	1721	1	I58902
26	106.5	2.4	2083	2	T42721
27	106	2.4	737	2	T16737
28	106	2.4	830	2	T37973
29	105.5	2.4	914	2	B96592

30	105.5	2.4	1272	2	C96637	hypothetical prote
31	105	2.4	1487	2	S15904	alpha-1 proteinase
32	104	2.4	1091	1	IUCHNL	neutral cell adhesi
33	104	2.4	2180	2	A47651	zinc-finger protei
34	104	2.4	2464	1	ORMSP1	microtubule-associ
35	103.5	2.4	558	2	T37567	probable NADPH cyt
36	103.5	2.4	727	2	G01792	transcription fact
37	103.5	2.4	818	2	S57078	probable protein K
38	103.5	2.4	866	2	T01171	GI/S transilition c
39	103.5	2.4	1163	2	I56126	lymphocyte function
40	103.5	2.4	4131	2	T21085	hypothetical prote
41	103	2.4	449	2	A53162	procollagen I C-pr
42	103	2.4	1565	2	AD2135	polyketide synthas
43	103	2.4	2347	1	TVHURS	kinase-related pro
44	102.5	2.3	810	1	A33380	interleukin-4 rece
45	102	2.3	418	2	T29737	hypothetical prote

ALIGNMENTS

RESULT 1
T08618
intrinsic factor-B12 receptor CUBILIN precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T08618
R:Moestrup, S.K.; Kozyraki, R.; Kristiansen, M.; Kayser, J.H.; Rasmussen, H.H.; Braut
J. Biol. Chem. 273, 5235-5242, 1998
A:Title: The intrinsic factor-vitamin B12 receptor and target of teratogenic antidiode.
A:Reference number: Z16459; MUID:98148073; PMID:9478979
A:Accession: T08618
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3623 <MOE>
A:Cross-references: UNIPROT:O70244; UNIPARC:UPI00000E776A; EMBL:AF022247; NID:93834379.
A:Gene: CUBILIN
C:Superfamily: Intrinsic factor-B12 receptor cubilin; EGF homology
C:Keywords: egg yolk; endocytosis; glycoprotein; intestine; kidney; peripheral membran
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-3623/Product: intrinsic factor-B12 receptor CUBILIN #status predicted <MAT>
F:133-164/Domain: EGF homology <EGF1>
F:436-467/Domain: EGF homology <EGF>
Query Match 3.1%; Score 137.5; DB 2; Length 3623;
Best Local Similarity 18.9%; Pred. No. 0.58;
Matches 151; Conservative 77; Mismatches 283; Indels 289; Gaps 34;
OY 70 IKSGERIVFTSSQSPENHFVIEIQKIDCKSGPCPFGEVQLQSTSLIPTLNRTFI 126
DB 2306 VSSREIRYLKFKHTGGSSYMGFKAKYASICGCTVS-GDSGVISIGVPTLPYANNVPCQ 2364
OY 127 MDVAHKSIGLELFSTIRLROIGRSGCPGCV-----HISGGLIDANVAVIGFCSNGT 182
DB 2365 WFIGLPEHYLTLSFEDPNL-QSSPG--CTKDYEIVEMHTSGRV-----LGRCGNST 2415
OY 183 VSRIMQGVGMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGE--GSAITLMSAN 239
DB 2416 PSSVDTSNNV-ASVKFTLDSVTASGRLQKSS-RQVC-----GGDLHPTGTTFSPN 2467
OY 240 YPEGPFDELMTWQFVPAHLRASVFLNENLSN-----C 274
DB 2468 YPNENPHARICEMWITVOEGRRIVLFTNLRLSTQPSGSEHLIVFNIGRNSPILQKLC 2527
OY 275 ERKKEVEVYIRGSTNPEVFKLEDKP-----GNMAGNFSLSQGDQDAQS 322
DB 2528 SRVAVTVEFKSSGNTMK-VVFTTGSRPYGGFTTASYSTEDAVCGGLPSPVSG--GNSS 2584
OY 323 PGI-----LRLQFOVLVQHPQNSNKIYVVDLSNERAMSLLIEPRPVKQSRKFPVG 373
DB 2585 PGVNGIRDYANLDCENTLISNPNRNSGISIYFLDELISHQDCT----- 2629

```

Qy 374 CFVCLSEKSSNLTGSGSKHKSFLCDDLRLMNVNVEKTSCTDHRVCOKRSYSLQVP 433
Db 2630 -FDVLEFFVGDADGSL-----I-EK-----FC-----SLAP 2654
Qy 434 SDILHP-----VELHDF----- 446
Db 2655 TAPLVIYPQWIFHFNVERVEYTGFIIEVSFTDCGIRTDNGVSSPNYPNLSAMTH 2714
Qy 447 -SMKLLVKKDLISLVLPQAQLOQHTHEKPCNTSSYLVASLIPQODLYFGFCGGSIK 505
Db 2715 CSMILKAEHGHTITLTFSDFLFAH---PCTSTSVTVRNGDSFGSPVIGRYCGQSVPR 2770
Qy 506 QIQVKN-1SVTLRTFASFQXOASRQGLTVSFIYFKE-----GVFTVPTDTSKYVL 559
Db 2771 PIQSSNQLIYFNT-----NNQOTRGFATWTNMGCGGTHSANGT-----I 2816
Qy 560 RTPNMDRLPLSLTSVSW-----NISVPRDQVACL-TFFKERSG----- 596
Db 2817 KSPHWPQTFEPENSRCSMTVTITHESKHWEISFDSNFRIPSSDQCONSFEVKWEGRLMINK 2876
Qy 597 -----VVCQGRAMITIQEORTRAEILFS----- 620
Db 2877 TLATSCGDVAPSPVITSGNIFTAVFQSEBEMAAQGFASFLSRGRTPTNTSGDIISPNF 2936
Qy 621 -----LDED-----VLKPSFH-----HSHFWNISNCSPTSGKQ 650
Db 2937 PKQYNNNNCTYLIDADQSLVILTFVSHLEDRSAITGTCHDGLHI-INKRNLSITPL 2995
Qy 651 LDILFSVTLTPRTVDLTVIL 670
Db 2996 VTICGSETLRPLTVDGPIVL 3015

```

RESULT 2

```

T33135
hypothetical protein C45G7.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33135
R:Dante, M.; Wamsley, P.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid C45G7.
A:Reference number: 221288
A:Accession: T33135
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-986 <DAN>
A:Cross-references: UNIPROT:O76355; UNIPARC:UPI000007ECAB; EMBL:AF067611; PIDN:AAC19183.
A:Experimental source: strain Bristol N2; clone C45G7
C:Genetics:
A:Gene: CESP:C45G7.5
A:Map position: 4
A:Introns: 373/3; 466/1; 524/3; 559/3; 722/2; 776/3; 839/2; 969/2
C:Superfamily: Caenorhabditis elegans hypothetical protein C45G7.5

```

Query Match 2.8%; Score 121.5; DB 2; Length 986;

Best Local Similarity 18.3%; Pred. No. 1.3; Indels 223; Gaps 32;

Matches 128; Conservative 101; Mismatches 243;

```

Qy 127 MDVXAKSIGLELQFSIRP-----LRQIGPGSCPDG 158
Db 387 FETKMMVACELAHMSVPREKRSQLECKLEKLLQALRENNKRLLYLRGMGHKT-PVK 445
Qy 159 VTHSGISGIDATVAVIGFCNGTYSRIKMGQGVKALHLFPFHRRNNSGFSIARRSSIK 218
Db 446 LAES-----STVENISKRETDTSVGRIQRSYTRL--LPYLKETPEWH----- 488
Qy 219 RLCTIESVFEGBGATLMSANY-----PEGFPED-----ELMTQFVVPALHLRASYFLNF 269
Db 489 QYMTIQDL---EKSDVESPNPIQNPQKFLFTEPRALLLTATNSGNEIYTSSTLKI 545
Qy 270 NLSNCKREERV-----EYIPGSTTWNEVFKLEDKQGNAGNFNISLQCCDDAOSP 323

```

```

Db 546 RVENLDNDEPPEFLPSALPIQVQPKNTSKPTAIGRLTARDADSPFIFHYLLPNCGTPSSD 605
Qy 324 GI-LRLQGVLYQHPO-NESNKIYVVDLSNERAMSLTEPRPVQSRKFVGCFCLESR 381
Db 606 NFNIDAEFGELVYEKVPSSKMTAEV-----CFIATQK 641
Qy 382 TC-SSNLTLTSGSKH---KISFLCDDLRLMNVNVEKTSCTDHRVCOKRSYSLQVPSDI 436
Db 642 NLDISEVFADNSKNFKKVKVEFGD-----SDVSVAANGKINNGFSGNGSISVDV 694
Qy 437 L---HLPEVL-----HDPKMLVPPKRLSLVLPQO-----KLOQHTHE 473
Db 695 LDRVEIPNMLCAAGADQYEOK-----SLNFPAYELGRDMSIPGAEVLEKNKGE 746
Qy 474 KPCNTSFYLVASA---IPSDLY---FGSCPGSSIQIQVKN-----ISVLTPTF 520
Db 747 -----LVANGKILDPQGVYTAELGRDGGGGVQQLHINRDKRLRYLSNRNEF 797
Qy 521 APSFXOASRQGLTVSFIYFKEGCVTPTDTSK---KYULRTPMMDRGLPSLTSVS 575
Db 798 GANLEK-----FRQILBAIGKDDQAKQQLIHPEPKADKRNSTWTSV- 842
Qy 576 WNISVPRDQVACLTFEKERSGVVCGTGAFMIIQEQRTABEILSLDEVLPKSFHHS 635
Db 843 -----CFYLTRONAILDENQASSISPSNGHISKL-----HHI 875
Qy 636 FMV-NISNCSPT-----TSKQDLDFSVTLTPRTVDLTVILAAVGGVLLLSALG 685
Db 876 FKQVNDMCAFKRPAATSESTSSNSDIPLN-----TLIIGVGLLIIMLLALL 925
Qy 686 LIICCVK---KKKTKN---GPAVGIYNNINTEMPROP 719
Db 926 IYVCCVSRYQRYLKKQTKDQLRCSSAGSYKSPNLIPEPPP 966

```

RESULT 3

```

T34022
zonadhesin - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
C:Accession: T34022
R:Hardy, D.M.; Gathers, D.L.
J. Biol. Chem. 270, 26025-26028, 1995
A:Title: A sperm membrane protein that binds in a species-specific manner to the egg ext
A:Reference number: 221464; MUID:96064658; PMID:7592795
A:Accession: T34022
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2476 <HNR>
A:Cross-references: UNIPROT:Q28983; UNIPARC:UPI000013C373; EMBL:U40024; NID:GI066465; PI
A:Experimental source: strain Meishan; testis
C:Genetics:
A:Gene: Zan
C:Function:
A:Description: may be involved in sperm adhesion to the zona pellucida

```

Query Match 2.7%; Score 118.5; DB 2; Length 2476;

Best Local Similarity 22.5%; Pred. No. 8.6; Indels 103; Gaps 18;

Matches 78; Conservative 45; Mismatches 120;

```

Qy 530 RQGLTVSFIYFKEGCVTPTDTSK-VYLRTPMMDRGLPSLTSVSNISVPRDQVACL 588
Db 172 RQGNKNTFI---QPAQPGISLNGGHYIFLETDFQAGQSFRLVSRPFCAP-AVICV 226
Qy 589 TFFKERSGVVCGTGAFMIIQEQRTABEIF-----SLDEVLK-PGFHHHSVW 638
Db 227 TFTYHMYGL-----GQGTKRLLLGSPAGSPSSLWERVVGQSPF-----WL 268
Qy 639 NISNCSPTSGKO-LDILFSVTLTPRTVDLTVILAAVGGVLLLSALGIICCVKKKKK 697
Db 269 NTSVTIPSGHQDPMLDIE-----AVRGNTAFVALGFVL----- 304
Qy 698 TNKGPAVGIYNNINTEMPROPKFKQKGRKNDNSHVAV-IEDTWYGHLLQDSSGSFLQ 756

```

Db 305 INHGTCRGPSETSVSTKCVAPRE---KPTVSEIYITPTKPMNH-----MKX 350
Qy 757 PEVDYRPFPGTMGVCPSPPITCSAPPAK-----LATEEP-----PRSPESSES 803
Db 351 PIVHTEKP---TVPTKEKPTIPTEKSTVPFKKPTVFKKPTLPREGPTVPAERPTTPEGPA 407
Qy 804 EF-----YTFSPNNNGDVSCKOTDP---LXKTOEPMEPAE 836
Db 408 VPKGPPTVLTWPTSTHTEKSTVTEKPIPLPTGKSTIPTEKPMVPTK 453

RESULT 4
558984
development protein tolkin (EC 3.4.24.-) - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C/Accession: S58984
R.Finelli, A.L.; Xie, T.; Bossele, C.A.; Blackman, R.K.; Padgett, R.W.
Genetics 141, 271-281, 1995
A/Title: The tolkin gene is a tolloid/BMP-1 homologue that is essential for Drosophila
A/Reference number: S58984; MUID:96042912; PMID:8536976
A/Accession: S58984
A/Molecule type: mRNA
A/Residues: 1-1464 <FIN>
A/Cross-references: UNIPROT:Q24132; UNIPARC:UPI00000820A1; EMBL:U34777; NID:g1002985; PI
A/Note: the authors did not translate the codon for residue 722
C/Genetics:
A/Gene: tolkin
A/Cross-references: FlyBase:FBgn0004885
C/Keywords: hydrolase; metalloproteinase; zinc
F:529-722/Domain: astacin homology <AST>
F:958-993/Domain: EGF homology <EGF>
F:1118-1153/Domain: EGF homology <EGF1>
F:614,618,624,673/binding site: zinc (His, His, His, Tyr) #status predicted
F:615/Active site: Glu #status predicted

Query Match 2.7%; Score 117.5; DB 2; Length 1464;
Best Local Similarity 19.3%; Pred. No. 4.6;
Matches 98; Conservative 63; Mismatches 161; Indels 185; Gaps 24;

Qy 105 PGCEVOLQSTSLPLTINRTFIMDVAKHSIGLELOFSIRLRQIGPSCGPDGYTHSIS 164
Db 1004 PNGTI-TSPSPFEMTPLKECIMEIYAPPKRISLNFTHPD-----EGRAHQOS 1052
Qy 165 --GRIDATVV-----RIGTFCSNGTVSRIKMQEGVKALHLPWFHPRNV--SGFSI 211
Db 1053 DCGYDSVTVYSKLGEMRLKRIGTFGSSSIPPTATSSN--ALRLFFHSDKSIQSGPAA 1109
Qy 212 ANRSSIKR-----LCITISVF-----EGEGS-----ATL 235
Db 1110 VFTTIDECVANNNGCOHBCRNTTISYICMCHNGVSMHENGHDKCEGEKYEISAPFGTI 1169
Qy 236 MSANYPEGPEDELTMTQFV--VPAHLRASVSFLNPL--SNCRKERVEVYIIPGATTNE 293
Db 1170 FSPNTPDSYPPNADCVMHFTTPGH-RIKLIPNEPDVESHQCTTDNVAVY----- 1219
Qy 294 VFLEKDPQGNNAAGNFNLISLQCDQDAQSPGIL-----RLOFOVLVQHPONESNKIYV 347
Db 1220 -----DGESESSVILGRFCGDKIPFI-----SSTSNQMWV 1251
Qy 348 DLSNERAMSLTEPPVYQSRKFVPGCFPLESRFCSSNULTLTSGSKHKSIFL-----C 401
Db 1252 -----LKTDKNKQKNGFASHSTACGGYLRATISQVQOQFYSHARFGNDY 1295
Qy 402 DDLRLMMNVKTIKCTDHRVYQ-----RKYSV-----LQVPSDLH----- 438
Db 1296 DD-----GMDCEWTIAAPNSYQVLFLFTPDIESSSEKCTPDYQVSDIDVYGQYPMKG 1351
Qy 439 -----LVELHDFSMKLL-----VFKDRLSLVLVPAQKLQDHT--EKDCNTSFSYL 483
Db 1352 QYCGNVLPQDINSMTSHLLVRFKTDGQSVMKGFASVYAVPNVPSGEYHSDHDEVENSYSE 1411

Qy 484 VASATPSQDLVYGRFCPGSGIKQIQVK 510
Db 1412 WVTPPP-----GSLKSYIE 1426

RESULT 5
A30788
mannose 6-phosphate receptor protein, cation-independent - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: A25908; A30788; S09404
R.Lobel, P.; Dahms, N.M.; Kornfeld, S.
J. Biol. Chem. 263, 2563-2570, 1988
A/Title: Cloning and sequence analysis of the cation-independent mannose 6-phosphate r
A/Reference number: A92706; MUID:88115411; PMID:2963004
A/Accession: A25908
A/Molecule type: mRNA
A/Residues: 1-2499 <LOB>
A/Cross-references: UNIPROT:P08169; UNIPARC:UPI000012F4B2; GB:J03527; NID:g162873; PIDN
R.Glickman, J.N.; Conibear, E.; Pearce, B.M.F.
EMBO J. 8, 1041-1047, 1989
A/Title: Specificity of binding of clathrin adaptors to signals on the mannose-6-phospi
A/Reference number: S09404; MUID:89305502; PMID:2545438
A/Contents: annotation: HA-II adaptor binding
C/Comment: This protein binds phosphorylated lysosomal enzymes and insulin-like growth
C/Superfamily: mannose 6-phosphate receptor, cation-independent; fibronectin type II r
C/Keywords: Golgi apparatus; membrane protein
F:1912-1951/Domain: fibronectin type II repeat homology <2F1>
F:2360-2363/Region: HA-II adaptor binding

Query Match 2.7%; Score 117; DB 1; Length 2499;
Best Local Similarity 19.8%; Pred. No. 11;
Matches 191; Conservative 103; Mismatches 343; Indels 326; Gaps 55;

Qy 1 MAGLNCGYSIALGLVLLGAARLPGAFAETALPRESNITVILKGPITLL---AKPC 56
Db 1710 MHGLACPRAGTAVCKRPVVG-----PIDIRVAGPPLINIANEVINFESSSTPC 1759
Qy 57 YIVISKRHITMLSIKSGRIVFTFSCQ-----SPE-----NHFVIEIQKNIDC---M 100
Db 1760 ---LADRFHNYTSL-----ITFHCKRGVSMGTBKLRTSCVDFPEMETVLCVPEVK 1809
Qy 101 SGCPCFGEVOLQPSLSLPLTINRTFIMDVAKH--SIGLELOFSIRLRQIGG--SCPDG 158
Db 1810 TDGCSLTBOQLYSPNLSSLSKSTPKVTRGPHTYSVG-----CTAAAGIDEGCKXDG 1862
Qy 159 VTHSISGRIDATTVVIGTFCSNGTVSRIKMQEGVKALHLPWFHPRNVSGFSIANRSSIX 218
Db 1863 AVCLLSGSKGASFGRLA-----SKLDYRHODEAVIISYANGDTCP 1903
Qy 219 -----RLCITESVFEGEASATLMSANYPEGPEDELTMTQFVPAHLRASVSFLNPLS 272
Db 1904 PETEDGEPVPPFVENGK-----SYEECVESBARLM-----CATTANYORD 1945
Qy 273 N-----CEKERREVEYIIGSTTNPVEVFKLEDKQFGNMAGNRNLSLQCDQDAQSGIIRL 328
Db 1946 HEWGFCKHSTSH-----RTSVIIFK-----CDEBAD----- 1971
Qy 329 QFOVLVQHPONESNKIYVVDLSNERAMSLTEPR-----PVQSRKFV----- 371
Db 1972 -----VGRPQ-----VSEVRGCEVTTEMKTKVVCPRKKECKVQKHRTYDALL 2017
Qy 372 --PGCF-----VCLF-----SRTCSSNLT-----TSGSKHKSIF-----LC 401
Db 2018 SLITGSMGFVHNGASYINLCQKIYKGPQDCSERASVCKKSTSGEVQYLGVHFTOKLIV 2077
Qy 402 DDLRLMMNVKTIKCTDHRVYQ-----RKYSV-----LQVPSDLH-----PVELHDSW 448
Db 2078 DD--RVITVYSKGHVCGNKTASAVIELTCAKTVGRPSFTFPDVSCTYH-----FSW 2128
Qy 449 KLVVFKDRLSLVLVPAQKLQ--QHTHEKPC--TFSYLVASATPSQDLVYGFSPCGSGIX 505
Db 2129 -----DSRAACAVKPGQEVQVQVNGITITPAGRSTSL-----GDIYFKRFSASGDVR 2174

Qy	506	-----QIOVKONISVTLKTPAPSFx-----	-----QZASRQGLTVSITPPFKKEG-	545
		: : : : : : : : : : : :	: : : : : : : : : : : :	
Db	2175	TNGDRYIYEIOL\$--\$IT-\$GSSPAC\$GASICOR\$KANDQHF\$SRKVGT\$NQTR\$YVVDGD	2230	
Qy	546	---VFT-----VTPDTSK\$VYLr-\$P\$MNDROGL\$P\$LTSV\$MNI\$SVPRDQVACLT\$PFKE	593	
		: : : : : : : : : : : :	: : : : : : : : : : : :	
Db	2231	L\$DVFTSSKCKGCKDKTSV\$STTFPHCDPL\$YKDG\$PE\$S-----	HE\$ADCOYL\$FSW	2281
Qy	594	R\$GVVCOTGR\$FMIQ\$EORTRA\$EELF\$SLED\$VLPK\$P\$FHNS\$FWVNIS\$NC\$PT\$GKO\$DL	653	
Db	2282	HT\$AVCPLGAGF-----	DEELADGD\$AEHK-----GL\$ERQ\$AVGAVL\$SL	2321
Qy	654	L\$F\$VTLT\$PRYDVL\$YIL\$A\$VGGVLL\$SALGLITC-----	YK\$K\$K\$K\$TKK\$PAVGIYN	708
Db	2322	L-----L\$VALT\$CLTL\$LL\$YK\$K\$R\$R\$EMV\$SR--L\$TNC\$CR\$SANV\$YK\$S\$K\$V\$K\$E\$B\$AD\$ENE	2375	
Qy	709	XNINTE--MP\$OPK\$K\$F\$OK\$R\$K\$ND\$SHV\$Y\$V--	I\$EDT\$--VY\$GHLO\$D\$SS\$P\$Q--PEVD	760
Db	2376	TE\$M\$T\$EEI\$Q\$P\$AP\$R\$K\$E\$G\$O\$E--NGH\$V\$AK\$S\$VR\$AD\$IT\$S\$AL\$IGD-EQ\$S\$E\$B\$E\$V\$LT\$LP\$V\$K	2432	
Qy	761	TYR\$P\$Q--GT\$MGVCP\$P\$P\$TIC\$SR\$AP\$TAK-----	L\$ATE\$P\$P\$P\$S\$P\$E\$S\$E\$P\$Y\$T\$F\$H\$N	812
Db	2433	VR\$P\$GR\$AR\$G\$AG--G\$P\$RL\$P\$R\$K\$AP\$P\$LP\$AR\$DR\$V\$GL\$V\$G\$E\$P\$AR\$R\$P\$PA\$AT\$PI\$ST\$F\$HD	2491	
Qy	813	NGD	815	
Db	2492	SDE	2494	

RESULT 6
T09456
intrinsic factor-B12 receptor Cubilin precursor - human
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T09456
R:Koyzraki, R.; Kristiansen, M.; Silahatoglu, A.; Hansen, C.; Jacobsen, C.; Tommerup, N.
Blood 91, 3593-3600, 1998
A:Title: The human intrinsic factor-vitamin B12 receptor, cubilin: Molecular characteri-
tion.
A:Reference number: Z16677; MUID:96241400; PMID:9572993
A:Accession: T09456
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3623 <KOZ>
A:Cross-references: UNIPROT:060494; UNIPARC:UPI0000070C4D; EMBL:AF034611; NID:g3929528;
C:Genetics:
A:Map position: 10p12
C:Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology
C:Keywords: receptor; vitamin B12 uptake
F:1-24/Domains: signal sequence #status predicted <SIG>
F:25-3623/Product: intrinsic factor-B12 receptor #status predicted <MAT>
F:436-467/Domains: EGF homology <EGF>

```

OY      249 LMTQPVVPAPHLRASVSFLPNLSNCRKEERVEYIIPGSTYPRVFKLED-----KQP-    302
Db      615 DCVAVITVSPDLVTFTEFTGLSLEHND-----DCKNDYLEIENDPLVYDPLL    661
OY      303 GNMAGNFNLSIQGCDQDAOSPG-ILRLQFOVLVQHPONESNKIYVVDLSNERMAMSLTIEP    361
Db      662 GKFTCTTSV-----PLQTTGGPPARLHFHSDSQISIQGCHITLTYLSPSRLRCGANTDP    715
OY      362 RPKVQSKRFVP---GCFVCLIESRTCSNLTLTGSSKHKISFLCDLTRLMMNVEKTI SCT    418
Db      716 ---EGELFPELISGPFT--HTRCQVYMMKQPOCEQIQINF-----THVE-----    754
OY      419 DHRICQKRSISLOVPSDLHLPVFLHDPFMSKLVLPKRLSLVLPQKLOQHTHEKPCNT    478
Db      755 -----LQQSSSSQNYIEVRD-----                          770
OY      479 SFSYLVASAIKPSODLYFGSFCPGGSIKQIOVKONISVTLTRFAPSPFOEAS--RQGLTVSF    537
Db      771 -----GETLLGKVCQNGTISHISITN-SWIRFKIDASVEKKSFRVIVQVAC    817
OY      538 IIPYKEGVPFVTPTDRTSKVYLRLTPNMDRGLPSLTYSWNIISVPRDQVACLTF--FKERS    595
Db      818 GDELITGSGV-----IRSPFPNNYVPGERTCRWTHIQQSQVILLNFTVFEIGS    865
OY      596 GVVCQT    601
Db      866 SAHCET    871

```

RESULT 7
 AH2011
 heterocyst glycolipid synthase [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AH2011
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchihara, T.; Tanaka, K.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference number: AB1807, MUID:21595285, PMID:11759840
 A:Accession: AH2011
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1263 <KUR>
 A:Cross-references: UNIPROT:Q8YWG7, UNIPARC:UP100000CE16B, GB:BA000019, PIDN:BA878012.1,
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: hglE

F125-3623/Product: intrinsic factor-B12 receptor #status predicted <MAT>	
F1436-467/Domain: EGF homology <EGF>	
Query Match	2.7% ; Score 117 ; DB 2 ; Length 3623 ;
Best Local Similarity	18.2% ; Pred. No. 20 ;
Matches 110 ; Conservative	77 ; Mismatches 229 ; Indels 190 ; Gaps 26 ;
QY	52 LAKPPIYIISKRHITMISKGERI--VFTSSCOSPEHHFVIEIOKNI-DCMSGCPPEPG- 107
DB	400 LSHPCL-----NGOCIDTVSGYFCCKDSCGWTGNCNTEINECTLSNPCLING 445
QY	108 -----EVLQPSLTLPLNRTF-----IMDYK 130
DB	446 TCVDGVDSPSCETRMLTGMALCOVPQVCGESLSGINSFYSRSPDVGYVHDVNCFWIK 505
QY	131 AHSIGLELOFSIPRLROIPEGSCCPDQVTHSISGRIDATVRIETFCNSGTVSRIKMOE 190
DB	506 TEMGKVLRIITLTFPLLESM---DNCPEHFLQVYDD-SSNAFQOLGRFCGSSLPHHLISSD 561
QY	191 GVKRALHLPWHPPRNVSGFSIANNRSSIKRLC--IIESVFEGBGATLMSANYPEGPEDE 248
DB	562 NA-LYFHLIYSEHLNNGRGFTVWMETO-OPEGCGILTGCV-----GSISPGPGNGVPPGR 614

```

Query March 2.6%; Score 113.5; DB 2; Length 1263;
Best Local Similarity 19.3%; Pred. No. 7.4;
Matches 170; Conservative 115; Mismatches 350; Indels 247; Gaps 42;

QY      TFSCQSENNHFVIE-----IQR-NIDCMGSPCPFG-----EVQLQPSSTLL 118
Db      TINTQENPKKIEESPFYINTERPMWIONGDIIRAGVSPFGCGTGVHVIVLEETPL 476
QY      119 P-----TINRTFINDVKAKHSIGLELQPSIRLQIQGPGSCPDGVTHSISGRIDAT 170
Db      477 PGGDRIHSTQSLILW-ADTPQQLSOKCATLEQLQSDNGEGEYNOLQNHNSKALLIPET 535
QY      171 VVRIGTCSNQTVARIRKMGEGVKMALHP-----WFHPRNV-----SGSINRRSSIRKLCT 222
Db      536 SARQGVATISLEEQKLLKAIKQLQIQPAFAEHHPQGIYYRQTGSLGK----- 587
QY      223 IESVFEEGSATL-MSANYPEGFPE-----DELTWQFVVPAAHLRAVSFLNFMLSN 273
Db      588 VVALFPQGGSYLWNGREISLNFPEIQAYQALDQLMA--QDNLP-----LSD 634
QY      274 CERKEERVEYYIPGSTTNPEVFKLEDKQPGNMAGNFMLSLQGCDDQAQSPGILRLQFOVL 333

```

```

Db      635 -----VFPIIP--AFSPETIKQSQQ-----LQRTENAPAIIGALSGLYKI 674
Qy      334 VQ-----HPONESNKIYVDLSNERAMSLTIERPKYQSKKFPGCCFVLESST 382
Db      675 LQKVGKPDVFAHGSEGLTALMAAGVFSDEDEYCYLIKAR--GQAMTIPOG--SHD 726
Qy      383 CSNNLT--TSGSKHKIKSFLCDDLTRLMM-----NVEKTIISCTDHRX 422
Db      727 CDGRTMLAVSGDVAKIKQIAGTKYQIANYNAPEQVULSGSKPEIALLEKULS----- 780
Qy      423 CQRKYSYIQ--VPSDILHLPVELHDPFSWKLVPKRLSLVLP--AQKLOQHTHEKPCNTS 479
Db      781 --KQGYTVPPLTVSAFAHTPFVRH-----ASQPFMALRLVTFNTPQIVR 823
Qy      480 ESYLVASAIPOSDLIFGSCFPGSGIKQIQVKONISVTLTAPSPFXQEASRQGL----- 533
Db      824 YANMTGNAYPTEATATIRQLLEAHLNAAVOFAQEIENLVAQGYCFVEIGPROILTNLVKQ 883
Qy      534 TVSFIPYFKEEGVFVTVPDTK--SKVYLK-----TPNMDRGLPSLTS 573
Db      884 TLGDRPHI---AIALNPERQKSDVQLKQGVYQLRWGLSLIEDDASP--RKLPAQKS 936
Qy      574 VSMNISVPRDQVACLTFFKERSGVWCOTGRAF-----MIOEQRTAEIIFSLDEVL 626
Db      937 KSKGSLI---KLMTATYISDK-----TKAEFEPAKPEPVQLASTPLDQVADAAATV- 986
Qy      627 PKPSTHHHSFWYNISNCSPTSGKQDLDFSVTLTPTPTV-DLTVILIAVGGVLLISALG 685
Db      987 -----VSDSPAQPLVASND--EIAITPVSWEEVSVL-----GEIYQSDYQ 1027
Qy      686 LIICCVKKKKKTKNKPAAVGINXNINTEMPPQPKFKQGRD---NDSHYAVIETDM 741
Db      1028 LLI-----SKDET-----TMNSTHNTVAVLFOQFYHQKEMQVHQYKTSQSQSQ 1076
Qy      742 VYGHLLQDSSGSLQPEVDYVRRPQGTGMV-----CPSPPTICSRAPTAKLATEPPP 795
Db      1077 AFLQLLQQQ--EVLIPQGSTSAIEGTVVVPAVVELPQTVASVVELKTVAVPVEBP 1134
Qy      796 RSPSESESEPTTFSSHNN--GVVSSQDTDIPILXTOEPNEPA 835
Db      1135 KVTAPVAPFAHISHANLEBRYSAPFPQPLVVTETASPA 1176

RESULT 8
Species: Mus musculus (house mouse)
C.Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C.Accession: S50065
R.Crocker, P.R.; Mucklow, S.; Bouckson, V.; McWilliam, A.; Willis, A.C.; Gordon, S.; Mill
EMBO J. 13, 4490-4503, 1994
A.Title: Sialoadhesin, a macrophage sialic acid binding receptor for haemopoietic cells
A.Reference number: S50065; MUID:95009950; PMID:7925291
A.Accession: S50065
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-1694 <CRO>
A.Cross-references: UNIPROT:O62230; UNIPARC:UPI0000028B74; EMBL:Z36293; NID:9557253; PIR
Query Match      2.6%; Score 113; DB 2; Length 1694;
Best Local Similarity 20.6%; Pred No. 13;
Matches 109; Conservative 64; Mismatches 191; Indels 166; Gaps 25;

Db      230 EGSATLMSANYPEGPEDELMTW-----QFVPAHLR-ASVSFLNPLNSCERKERV 281
Qy      903 EGOAVVLSCQVPTGVSEGTYSWYQDGRPLQESTSTLITAIASLRQAGAYHCQADAPT 962
Db      282 EX-----YIPGSTTNPEVFKLEDKPQNMAGNPN-----LS 312
Qy      963 AIASLAAPVSLHVSSTPRHVTLSALLSTDPERLGHVCSVQSDPPAQQLFHRNRLVAST 1022
Qy      313 LQGCQDQAS-----PGILRLQFVVLVQHPQNESNKIYVDLSN----- 351

```

```

Db      1023 LQGADELAGSNRLHVTVLPNELRLQ-----IHPPELEDGCTVTCESANTLQGAASAAPRD 1078
Qy      332 ERAMSLTIERPKVQSKRFPVCCFVLESRTSSNLTLTSGSKHKISFLCDDLTLMM--- 408
Db      1079 AQAVRVTWPNATVQEGQV-----NLTCILWSTHQ-----SLSYTWYKG 1119
Qy      409 -----MNVEKTIISCTDHRXC-----QKRSYSLOVPSDILHLPVELHDPFSWKL 451
Db      1120 GQQLLGARSITLPSVKVLDATSYR--CGVGLRGHAPHLSPVTLVDVLAHPRNLK-LTYLLE 1177
Qy      452 VPKRLSLVLV-----PAQKLOQHTHEKPCNTSFSYLVA---AIP----- 489
Db      1178 TQGRQLALVLCVDSRPPAQLTSHGDQ-----LVASSTEASVPTLRLLELDQDPP 1228
Qy      490 -SQDLY-FGSCFPGS-----IKQIQVKONISVTLTAPSPFXQEASRQGLTVSFT 539
Db      1229 SNEGLYSCSAHSPFLKANTSLLELLEGVAVKKNPQSGVDEGEPVTVTCEDPAALSSALVA 1288
Qy      540 YFK-----EEG-----VTVTPDTKSKVYLTRPMNDRGLPSLTSWNI-SVPRQVACL 588
Db      1289 WFHNGHMLQEGPASSLQPLVTTTRAHAGAYFCVHDQTGTRSRPASLQILVAPRD-AVL 1346
Qy      589 TFFKERSGVWCOTGRAPMIQ---EQRTAEIIFSLDEVLPKPSPHHHS 635
Db      1347 SSFRD-----SRTLMVVIQCTVDSERPPEAVVLSHNGKVLAAHSHRHS 1390

RESULT 9
Species: Homo sapiens (man)
C.Date: 13-Sep-1991 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C.Accession: A36688; A34210; A36613; A30187
R.Mikol, D.D.; Alexakos, M.J.; Bayley, C.A.; Lemons, R.S.; Le Beau, M.M.; Stefansson, K.
J. Cell Biol. 111, 2673-2679, 1990
A.Title: Structure and chromosomal localization of the gene for the oligodendrocyte-my
A.Reference number: A36688; MUID:91115958; PMID:2277079
A.Accession: A36688
A.Molecule type: DNA
A.Residues: 1-440 <MI>
A.Cross-references: UNIPROT:P23515; UNIPARC:UPI0000130CE1; GB:X57436; NID:g23105; PIDN
R.Mikol, D.D.; Gulcher, J.R.; Stefansson, K.
J. Cell Biol. 110, 471-479, 1990
A.Title: The oligodendrocyte-myelin glycoprotein belongs to a distinct family of protein
A.Reference number: A34210; MUID:90130636; PMID:1688857
A.Accession: A34210
A.Molecule type: DNA
A.Residues: 8-20, 'X', 22-440 <MI2>
A.Cross-references: UNIPARC:UPI000016AE0D; GB:X51694; NID:g35144; PIDN:CA35991.1; PID
R.Vibkochil, D.; Cawthon, R.; O'Connell, P.; Xu, G.; Stevens, J.; Culver, M.; Carey, J
Mol. Cell. Biol. 11, 906-912, 1991
A.Title: The gene encoding the oligodendrocyte-myelin glycoprotein is embedded within t
A.Reference number: A39613; MUID:91117257; PMID:1899288
A.Accession: A39613
A.Molecule type: mRNA
A.Residues: 1-20, 'X', 22-440 <VI>
A.Cross-references: UNIPARC:UPI0000163BC8; GB:M63623; NID:g189385; PIDN:AA559970.1; PIR
J. Cell Biol. 106, 1273-1279, 1988
A.Title: A phosphatidylinositol-linked peanut agglutinin-binding glycoprotein in centra
A.Reference number: A30187; MUID:88198371; PMID:3283151
A.Accession: A30187
A.Molecule type: Protein
A.Residues: 25-29, 'H', 31, 'H', 33-44, 'D', 46-50, 'P', 52 <MI3>
A.Cross-references: UNIPARC:UPI000017C2D8
C.Genetics:
A.Gene: GDB:OMG; OMGP
A.Cross-references: GDB:127563; OMIM:164345
A.Map position: 17q11.2-17q11.2
A.Introns: #status absent
A.Note: this gene lies within an intron of GDB:NFI on the opposite strand
C.Keywords: blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinositol linka
P.1-24/Domains: signal sequence #status predicted <SIG>

```


Db 481 EDDIISIIDNKAYHARVSTPIATQTLHAL-----FSTFLSNTYPCATQANTGTAMW 536
Qy 455 DRLSLVLVPAQKLOQHTHEKPCNTSPSYLVASAISS-----ODLYFGSFCGGSIKQ 506
Db 537 NDLSYLVNPLKSSVH---KPAIVSNSSRFVSVDYYPYLKVDIDQFRSIFFSVNDH 593
Qy 507 IOYKNISVTLR---TFAPSPYQF-ASROGLVSIPIPFKEGVTVP-----DTKSV 557
Db 594 LQVPEKL-VLLRYSSVWCPCPNKQKPSMAFVTLNHYLYNLISGFSYLRRIIDLIDISIE 652
Qy 558 YLRTP-----NMDRGLPSLTYSWNISY-----PRDQVA 586
Db 653 YDKSKKHYSRMLNHQRG-----DGLRFNMSVFTTRRAVASKLOFLINKAMHITKGEKE 708
Qy 587 CLTFPKERSGVVCGTGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHH 634
Db 709 VLEIFQELDEIEIENEK---IIKDLISESH-YSKDYDYLKSTYDHH 752

RESULT 12

J01674

protein kinase TMK1 (BC 2.7.1.-), receptor type precursor - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Dec-2004

C/Accession: J01674

R/Chang, C.; Schaller, G.E.; Patterson, S.E.; Kwok, S.F.; Meyerowitz, E.M.; Bleecker, A.

Plant Cell 4, 1263-1271, 1992

A/Title: The TMK1 gene from Arabidopsis codes for a protein with structural and biochemical

A/Reference number: J01674; MUID:93076110; PMID:1332795

A/Accession: J01674

A/Molecule type: DNA

A/Residues: 1-942 <CH>

A/Cross-references: UNIPROT:P43398; UNIPARC:UPI0000048404; GB:I00670; NID:G166887; PIDN:

C/Superfamily: Receptor-like protein kinase; leucine-rich alpha-2-glycoprotein repeat hc

C/Keywords: ATP; autophosphorylation; glycoprotein; phosphotransferase; receptor; serine

F/1-22/Domain: signal sequence #status predicted <SIG>

F/23-942/Product: protein kinase TMK1 #status predicted <MAT>

F/65-88/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F/89-111/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F/112-135/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F/136-160/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F/161-186/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F/187-209/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F/210-232/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F/233-255/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>

F/256-279/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>

F/280-299/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F/300-323/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F/324-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>

F/363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F/387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>

F/411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>

F/440-503/Domain: transmembrane #status predicted <TM>

F/566-872/Domain: protein kinase homology <KIN>

F/559-602/Region: protein kinase ATP-binding motif

F/86-99/158,161,171,353,533,587/Binding site: carbohydrate (Asn) (covalent) #status pred

F/616,634,717,719/Active site: Lys, Gln, Asp, Lys #status predicted

Query Match 2.5%; Score 111.5; DB 1; Length 942;

Best Local Similarity 19.4%; Pred. No. 6.7; Indels 233; Gaps 29;

Matches 128; Conservative 82; Mismatches 216; Indels 233; Gaps 29;

Qy 117 LPLPTLRTFTWDVKAHNSIGLEQFSIPRLRQIGSGSCPDGVTSHI-SGRIDATVVRIG 175

Db 16 LLTSLSKA---DSQGLSAMLSTKSLNPPSSFGMSDDPCMKWHIVCTGRKATRIQIG 72

Qy 176 TFCNGCYSTRKIMQGVMAHLPMFHRPNVSG--FSJANSSIKRLCTIISVBEG--- 229

Db 73 HSGLGQTLSPDLNLSELERLEQW--NNISGVPVPSISGLASIQVLMNNNPDSPISD 129

Qy 230 --EGSATLMSANYPEGFPDELTMTQFVVPALRLASVFLFNLSNCRKEKREVEYYIPG 287

Db 130 VFQGLTSLQSV-----IDNPFKSMW--IPESLNAASLQNFSAASA-----NVSGSLPG 178
Qy 288 STNPEVFKLEDKQPG-----NMAGNFNLSIQGCDQDQAGPGLRLQFQVLYQHPO 338
Db 179 -----FLGDEFPFGSILHLAFNNLEGEPLMSLAGS-----QVOSLMLNGQ 219
Qy 339 NESKIIYVDSLNSERASMLTEPRPVKQSRKFVGGCFPLESRICSSULTLTSGKHS 398
Db 220 KLTDITV--LQNNWTKL-----EYWLHNSKRFSGPLPDFSGIKEL-- 257
Qy 399 FLCDLTRLMMNVKSTICTDHRVCQKRSYSLQVPSDILH-----PYEL 443
Db 258 -----ESLSLRN-----SFGCPVPSLISLSLKVVNLTNNHLOGPVPV 297
Qy 444 HDFSMLKLVPRK-----LSLVLV-----PAQKLOQHTHEKPCNTSPSY 482
Db 298 FKSSVYDLDDKDSNFCILSPGECQPRVKSILLIASSFDYPRLAESWKGNDPCTN----- 353
Qy 483 LVASAISSQDLYPSFCPGGSIKQIQYKONISVTLRTFAPSFQDASRQGLTV-----S 536
Db 354 -----WIGIACSNGNITVLSL-EKHELT-GTISPEFGAISKLRITLGINMLTG 400
Qy 537 FIPYFKEGVTVTPDRK-----SKYVLRTPNMDRGLPSLTYSWNISVPRDQVACLT 590
Db 401 MIPQ-----ELTTPNLKTLDVSSNKLFGKVPGF----- 429
Qy 591 FKERSGVVCGTGRAFMIIQEQRTAEIIFSLDEVDLPKPSFHHHSFWNINSGSPTSG-- 648
Db 430 ---RSNVVV-----NTGNPDIGDKKSLSP-----GSSSPSGSGS 463
Qy 649 -----KQDLREVTLTRVVDLTVLIIAAGGVLLISALGILICVKKKKKKTNGK 701
Db 464 SGINGDORCKMKSSTPIG-----IIVGVLGLLSTFLIGLIVFCWYKQKQKRFSG 515

RESULT 13

T24315

hypothetical protein T01G9.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T24315

R/Leemard, N.

submitted to the EMBL Data Library, July 1996

A/Reference number: Z19874

A/Accession: T24315

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-603 <MT>

A/Cross-references: UNIPROT:Q22075; UNIPARC:UPI00000762F8; EMBL:Z75713; PIDN: CAB00050.1

A/Experimental source: clone T01G9

A/Genetics:

A:Gene: CESP:T01G9.3

A:Map position: 1

A:Introns: 68/3; 346/1; 407/3; 450/3; 486/3

Query Match 2.5%; Score 111; DB 2; Length 603;

Best Local Similarity 19.7%; Pred. No. 3.8;

Matches 132; Conservative 79; Mismatches 217; Indels 242; Gaps 32;

Qy 266 FLNF-----NLNCRKEKREVEYIPGSTTNEVF-----KLEDK 300

Db 9 FLSLVLVLCIALPSSCPNLCECDNDSSWSVCKAIINNTYIAEILINQLPLTSLRHQ 68

Qy 301 QPGNMAG-----NFN-----LSLQCCDDQDQAGPGLRLQFQVLYQHPON-----E 340

Db 69 PPSRIGSNKLRNMNDINRRAQLRVLRILNCOIPAMRSIRLPSLEVLDHSHNNIEHATM 128

Qy 341 SN-----KIYVDSLNSERASMLTEPRPVKQSRKFVGGCFVLES-RTGS-SNLTLTSGS 393

Db 129 SNFGQKMLKVLVLDSSNH--LNILP-----TGVEFYRALRSLSLSNNTTSLDLS 175

Qy 394 KHKISFL-----CODLTRLMMNVKSTICTDHRV---CQKRS-VSLQVPSD 435

Db 176 TNLKGLNSLRVLRDRPIPIEHINELFTDY-----SQDLDELXNHCNLSISYSLADR- 230
Qy 436 ILHLPEVLHDSKLLVKKDL-----SLVLPVPOKLOQTHREKPCNTSPSYLVASALP 489
Db 231 -----IPQLRQLGIGNNLKVPTELRS-----LPQLSVLDLSHNS 267
Qy 490 SODLYFGSPGSGISKOIQVKONI-----SVTLRTPAFSP----- 524
Db 268 IQEIRACAFPC-NTNISKLDLSHNLGISKDSPFNEDAPRTMPLRLDLSPNMDPDSKW 326
Qy 525 ---XQASRQGLTVSFIPYKEGCVTVTPDKSKVYLRTD-NMDRLP----- 569
Db 327 LGMAQEEITSLISNGFLNFEESWTYTL-----KSLIHLELAYNHIKFIPLVQLPARYHL 382
Qy 570 -----SLTSSVNISVPRDOVACLTFFKERSGVNCTGAPMIDQO----- 611
Db 383 ISLNIISGNETYLPNINITLLPNVKTFTDITANRFHTFSHTDLAFINNEQVYVDGNPMD 442
Qy 612 -----RTRAEIPSLDEDLVLPKPSF-----HHHSFWNINSCSPTSKOLDLLF 655
Db 443 SCATGLOVHRMDRVRAMHILNYDVNRCATPSLVESHSLATTDVNDCAVLFGAAYGL-- 500
Qy 656 SVTLRTPVDTLVILIAAVGGVLLLSALGLITC-CVKKKKKTKNKPAVGIYNNINITE 714
Db 501 ---TQISEMILIL-----AGVLLPALLMLITLIGIYFLRROYKG-----SYVTR 543
Qy 715 MPQPKKKQKGRKNDSHVYAVIEDTMYGHLLODSSGSF---LQPEVDTY---RPPQG 767
Db 544 HSRITL-----TMANTHSCSSSTMDTHGRLPPLPPDFLVSTETFXA 584
Qy 768 TMGVCPPSP 777
Db 585 TPRLPPAP 594

RESULT 14

T19212
hypothetical protein C12C8.3 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C.Accession: T19212
R.Barlow, K.
submitted to the EMBL Data Library, November 1996
A.Reference number: Z19091
A.Accession: T19212
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-1185 <WIL>
A.Cross-references: UNIPARC:UPI000017B772; EMBL:Z81467; PIDN:CA803872.1; GSPDB:GN00019;
A.Experimental source: clone C12C8
A.Genetics:
A.Gene: CESP:C12C8.3
A.Map position: 1
A.Introns: 46/1, 133/2, 150/1, 197/1, 287/3, 378/2, 419/3, 449/1, 553/3, 580/2, 704/2, 7

Query Match 2.5%; Score 111; DB 2; Length 1185;

Best Local Similarity 19.0%; Pred. No. 10; Mismatches 312; Indels 346; Gaps 48;

Matches 181; Conservative 114; Mismatches 312; Indels 346; Gaps 48;

Qy 17 ILGAARLPGRGAFAFIAPRESNITVLIKLGPTTL-LAKPCYIVISKRIITMLSIKSGR 75
Db 310 LMASPDVPR--HSATIAPRNS-----MCSLPRLQLATP-----MSGSQD- 348
Qy 76 IVFTSCSGPENHFVIEIKNIDCMGCPPEBVOLOPST-----LLPTLNRTFIWDV 129
Db 349 ---TPIPSPLGSOPQOO-----PMGPIDQCGESKISFAYCMQCEALCICV 395
Qy 130 KAHKIGLELQPSIRLRQ-----IGPESCDGVTSHISGRIDATVVRIGFCGN 180
Db 396 QHQRVATKQHAPELQOLMTLMSRAVQPOOA---QYTTQVNGSVROALSGVGFESG 453
Qy 181 GTVSRIKMGEGVMALHLPWFPRNVSGFSIANRSSIKR---LC-----IISVEGEGS 232

Db 454 -----HVSQVENDSIGSGSSPSRSSVCGTHDSVIIIGICENCPH 492
Qy 223 ATMS-----ANYP-----EGFPDELTWQFVVPALHRAVSFLNPLLS 272
Db 493 SVLLCAICVAOHPEGRVQPLGDIRAVAVGEVNVSSQLQWQ----- 533
Qy 273 NCERKEERVEYIIGSTTN-----PEVFKLEDQPG-----NNAGFNLT- 311
Db 534 -CEKGTGDITKQIIGIYVNTAATANEIRAAFDTHVNLAEERKELKLVETVYKOLKLSVL 592
Qy 312 ---SLQCGDQAO-----SPGILRLOFQVLYQ-----HPONESNK 343
Db 593 ISQAESLQSKOIIDQALQIATKLMSSDCDEMYLRQVFXFKLACQMGNEGTEPPNN--NI 650
Qy 344 IYVVDLS---HERAMSLIETPRP---VKQSKRFVPGCVLESRTCSNLLTSGSKHKI 397
Db 651 LNVMLACQVEDRLKFTAPQDGLLNKARQGN---IESGPCANSSIVGDSFKKI 705
Qy 398 SFLCDLTRLMMNVEKTICTDHR-----CQKRSYSLOVPSD 435
Db 706 -----INNEKMGMS--RYSNHFISQFELNCFIFQONQSEVCGKNIAIRERQT 752
Qy 436 ILHLPEVLHDSKLL-----VPKRLSLVLPQAKLOQTHREKPCNTSPSYLVASALPSQ 491
Db 753 VIY--VQLRDACGDLSSSIATOPTSQALLPHQ--EPHSH-----LEQAMPTS 797
Qy 492 DLYRGSFPGSGISKOIQVKONISVTLRTPAFSPFOQASRQGLTVSFIPYKEGCVFTTP 551
Db 798 DVQAFVISPDSSTVEV-----TMTD-----RENGIVALSTYPSIEGTYLNTI 839
Qy 552 DTKSKVYLRTNPMNDRLPSLTSSVSNISVPRDOVAC---LTFPKERSG-----VVC 599
Db 840 LVKG-----TP--ISGCTTMDIRGRNY--DEIAAKPILTFGSGSDDELCPWIC 890
Qy 600 QTGAEMTIOEORTAEIIFSLDEDLVLPKPSFHHHSFWNINSCSPISGK---QLDLFS 656
Db 891 VDQGRVIVADRSMNRVQIFDKDGNFISK-----FGTSGNPPQGFDRPAG 935
Qy 657 VTLRTPVDTLVILIAAVG-----GVLLLSALGLIICVKKKKKTKNKPAVGIY 707
Db 936 IT---TNSLNNIVADKONHRVOVFENGMFL-----KFGDRGRAVGIF 977
Qy 708 NXNINTEMPROPKKQGRKNDSHVYAVIEDTMYGHLLODSSGSFLQ---PEVDTYR 763
Db 978 N-----YPMGVATN--SHNALAVSDTNRHRVQITPQOQFVRKXGPDSEAYFFK 1023
Qy 764 PFOGTWVC-PSPPTTSRAPYAKLATBEPSPSPSESPYTFGHPNNGD 815
Db 1024 NLDSPFGICVLPDGLITIDFNNHRLAV-----LSPNMSEMKVYGSBGDD 1070

RESULT 15

T42205
breast cancer susceptibility protein BRCA2 - mouse

C.Species: Mus musculus (house mouse)

C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 15-Mar-2004

C.Accession: T42205
R.McCallister, K.A.; Haugen-Striano, A.; Hagevik, S.; Colline, N.K.; Brownlee, H.; Futreal

submitted to the EMBL Data Library, February 1997
A.Description: Characterization of the mouse and rat homologs of the BRCA2 breast cancer

A.Reference number: Z22073

A.Accession: T42205
A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA
A.Residues: 1-3329 <MCA>

A.Cross-references: UNIPARC:UPI000161CCC; EMBL:U89652; NID:G2443438; PID:G2443439; PIDN

A.Experimental source: strain CD1, 129Sv; ICR Swiss
A.Genetics:
A.Gene: BRCA2

C.Superfamily: DNA recombination repair protein, BRCA2 type

Query Match 2.5%; Score 111; DB 2; Length 3329;

Best Local Similarity 18.8%; Pred. No. 49;

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 22:12:25 ; Search time 119 Seconds
(without alignments)
4956.485 Million cell updates/sec

Title: US-10-781-564-1
Perfect score: 4377
Sequence: 1 MAGNCGVSIALGLVLLGA.....SSKTDIDPLKTFQEMEPAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 05.80:*
1: uniprot_sprot:*
2: uniprot_trcembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4375	99.9	836	2	Q9H5V8_HUMAN
2	4374	99.9	836	2	Q96QU7_HUMAN
3	3629	82.9	697	2	Q6HT71_HUMAN
4	3603	82.3	691	2	Q6U9Y2_HUMAN
5	3574.5	81.7	833	2	Q5U462_MOUSE
6	3395	77.6	649	2	Q9H8C2_HUMAN
7	2041	46.6	392	2	Q9H676_HUMAN
8	1849.5	42.3	470	2	Q8B2D9_MOUSE
9	1779	40.6	343	2	Q8WU91_HUMAN
10	1579.5	36.1	384	2	Q921M9_MOUSE
11	828	18.9	867	2	Q4T223_TETNG
12	824	18.8	825	2	Q4T7U3_TETNG
13	630	14.4	160	2	Q810U0_MOUSE
14	610	13.9	793	2	Q58EM7_BRARE
15	141	3.2	746	2	Q01654_HALRO
16	140.5	3.2	888	2	Q51F10_ENTHI
17	140	3.2	887	2	Q6ONK4_CAEBR
18	139	3.2	746	2	Q81AD8_HALRO
19	137.5	3.1	719	2	Q9PVY2_TRISC
20	137.5	3.1	3623	2	Q70244_RAT
21	131.5	3.0	2796	1	CSMD3_MOUSE
22	127	2.9	595	1	SIGL1_HUMAN
23	126	2.9	832	2	Q9JLB4_MOUSE
24	125.5	2.9	508	2	Q18530_CAEBL
25	123.5	2.8	3620	2	Q9UTU3_CANPA
26	123.5	2.8	4060	2	Q6T3A5_MOUSE
27	122.5	2.8	421	2	Q7YFC4_CRIYP
28	121.5	2.8	2302	2	Q5F4U0_CAEBL
29	121	2.8	766	2	Q618F0_CAEBR
30	121	2.8	1242	2	Q4X2J9_PLACH
31	120.5	2.8	3570	2	Q7G737_ANOGA

32	120	2.7	567	2	Q61DZ1_ANOGA	Q61d21 anopheles g
33	120	2.7	1516	2	Q504M1_XENTR	Q504m1 xenopus tro
34	119	2.7	597	1	SIGL1_PANTR	Q951h0 pan troglod
35	119	2.7	847	2	Q82J97_STRAW	Q82j97 streptomyc
36	119	2.7	941	2	Q4RY40_TETNG	Q4ry40 tetraodon n
37	119	2.7	2122	2	Q54R92_DICDI	Q54r92 dictyosteli
38	118.5	2.7	407	2	Q9D2J4_MOUSE	Q9d2j4 mus musculu
39	118.5	2.7	2476	1	ZAN_PIG	Q28983 mus scrofa
40	118.5	2.7	4059	2	Q8C1S7_MOUSE	Q8c1s7 mus musculu
41	118	2.7	944	2	Q6CSES_KLULA	Q6cses kluyveromyc
42	118	2.7	2520	2	Q6HFE0_BACHK	Q6hfe0 bacillus th
43	118	2.7	2766	1	PDZK3_RAT	Q94ZRS rattus norv
44	117.5	2.7	819	2	Q747S7_GROSL	Q747s2 geobacter s
45	117.5	2.7	1424	2	Q91TRS_ARATH	Q91trs arabidopsis

ALIGNMENTS

RESULT 1	Q9H5V8_HUMAN	PREDIMINARY;	PRT;	836 AA.
ID	Q9H5V8; Q9H5V8; Q9H5V8; Q9H5V8;			
AC	Q9H5V8;			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-FEB-2005 (TREMBLrel. 29, Last annotation update)			
DE	Hypothetical protein FLJ22969 (NCSG135).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	NUCLEOTIDE SEQUENCE.			
RA	Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,			
RA	Suzuki Y., Ohtsuka M., Nishi T., Shibahara T., Tanaka T.,			
RA	Nakamura Y., Isono T., Sugano S.,			
RL	Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.			
RN	(2)			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=22547370; PubMed=12660814; DOI=10.1038/sj.onc.1206220;			
RA	Hooper J.D., Zijlstra A., Aimes R.T., Liang H., Claassen G.F.,			
RA	Tatin D., Teeta U.E., Quigley J.P.,			
RT	"Subtractive immunization using highly metastatic human tumor cells			
RT	glycoprotein antigen."			
RL	Oncogene 22:1783-1794(2003).			
DR	EMBL; AK026632; BAB15511.1; mRNA.			
DR	EMBL; AF468010; AAO33397.1; mRNA.			
SQ	SEQUENCE 836 AA; 92874 MW; 98980475C3E5C4C8 CRC64;			
Query Match	99.9%; Score 4375; DB 2; Length 836;			
Best Local Similarity	99.6%; Pred. No. 3.2e-314;			
Matches	833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
QY	1 MAGNCGVSIALGLVLLGAARLPARGAEFIALPRESNTIVLIKLGPTLLAKRCYIVI			
DB	1 MAGNCGVSIALGLVLLGAARLPARGAEFIALPRESNTIVLIKLGPTLLAKRCYIVI			
QY	61 SKRIITMISTSGRIVFTSCQSPENHFVIEIQKIDMGCPFPFGVQLQPSLSLPT			
DB	61 SKRIITMISTSGRIVFTSCQSPENHFVIEIQKIDMGCPFPFGVQLQPSLSLPT			
QY	121 LNRFTMDVKAKHSIGLEQSFIRLRQIGGESCPGVTHSIGRIDATVVRIGTFCSN			
DB	121 LNRFTMDVKAKHSIGLEQSFIRLRQIGGESCPGVTHSIGRIDATVVRIGTFCSN			
QY	181 GTVSRIRKQSGVKALHPWFHPNVGFSIANSSTIKRLCTISVFEGBGSATLMSANY			
DB	181 GTVSRIRKQSGVKALHPWFHPNVGFSIANSSTIKRLCTISVFEGBGSATLMSANY			
QY	241 PEGPPEELMFWGVVAHLPASVFLNPNLSNCRKEERVEYVYIPGSTNPEVFXLEDK			
DB	241 PEGPPEELMFWGVVAHLPASVFLNPNLSNCRKEERVEYVYIPGSTNPEVFXLEDK			

```
Db 241 PEGFEDELMTMOWFVPAHLRASVFLNFNLSNCRKERVEYIIPGSTTNPVEFKLBDK 300
Qy 301 QPGNMGAFNLSLQCCDDAOSPGILRLOFVLYOHPONESNKIYVVDLSNBRASLTIE 360
Db 301 QPGNMGAFNLSLQCCDDAOSPGILRLOFVLYOHPONESNKIYVVDLSNBRASLTIE 360
Qy 361 PRPVAKSRKFVPGCFVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Db 361 PRPVAKSRKFVPGCFVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Qy 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVLPKXRLSLVLPKQLQOHTHEKPCNTSF 480
Db 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVLPKXRLSLVLPKQLQOHTHEKPCNTSF 480
Qy 421 SYLVASAIIPSDLYFGSFCPGGSIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFLPY 540
Db 421 SYLVASAIIPSDLYFGSFCPGGSIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFLPY 540
Qy 541 FKEBGFVTPTDTSKAVLRTPNMDRGLPSLTSSVMNISVPRDOVACLTFFKERSGVVQ 600
Db 541 FKEBGFVTPTDTSKAVLRTPNMDRGLPSLTSSVMNISVPRDOVACLTFFKERSGVVQ 600
Qy 601 PRPVAKSRKFVPGCFVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Db 601 PRPVAKSRKFVPGCFVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Qy 601 RYCQKRSYSLOVPSDILHLPELHDFSWKLVLPKXRLSLVLPKQLQOHTHEKPCNTSF 480
Db 601 RYCQKRSYSLOVPSDILHLPELHDFSWKLVLPKXRLSLVLPKQLQOHTHEKPCNTSF 480
Qy 661 PRVTLVTLIAVGGVLLSALGLITCCVKKKKTKNKGPVAGIYXNNINTEMPROPK 720
Db 661 PRVTLVTLIAVGGVLLSALGLITCCVKKKKTKNKGPVAGIYXNNINTEMPROPK 720
Qy 721 KFOGRKONDSHVAVIEDTMYGHLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
Db 721 KFOGRKONDSHVAVIEDTMYGHLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
Qy 721 KFOGRKONDSHVAVIEDTMYGHLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
Db 721 KFOGRKONDSHVAVIEDTMYGHLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
Qy 781 SRAPPAKLATEBPSPSPSESESEPTFSHPNNGDVSSKDTDIPILXTOEPMEPAE 836
Db 781 SRAPPAKLATEBPSPSPSESESEPTFSHPNNGDVSSKDTDIPILXTOEPMEPAE 836
```

RESULT 2

```
Q96GUT_HUMAN PRELIMINARY; PRT; 836 AA.
ID Q96GUT_HUMAN PRELIMINARY; PRT; 836 AA.
AC Q96GUT;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CUB domain containing protein 1.
GN Name=CDCP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RN (1)
RX NUCLEOTIDE SEQUENCE.
RP MEDLINE=2159860; PubMed=146621; DOI=10.1038/sj.onc.1204566;
RA Scherl-Mostaguer M., Sommergruber W., Abseher R., Hauptmann R.,
RT Ambros P., Schweifer N.,
RT "Identification of a novel gene, CDCP1, overexpressed in human
RT colorectal cancer.";
RL Oncogene 20:4402-4408 (2001).
DR EMBL, AY026461; AAK02058.1; -; mRNA.
DR Ensembl, ENSG00000163814; Homo sapiens.
SQ SEQUENCE 836 AA, 92873 MW, PBA4D2DBBD35C519 CRC64;
```

Query Match 99.9%; Score 4374; DB 2; Length 836;
Best Local Similarity 99.6%; Pred. No. 3.7e-314;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAGNCGVIALGVLLGAARLPRAAEFAELPRESITVILKIGTTLAKPCYIV 60
Db 1 MAGNCGVIALGVLLGAARLPRAAEFAELPRESITVILKIGTTLAKPCYIV 60

```
Qy 61 SKRHITMLSIKSERIVTFSCQSEBHFVIEIQNIDCMGSCPCFGEVQLQPSLTP 120
Db 61 SKRHITMLSIKSERIVTFSCQSEBHFVIEIQNIDCMGSCPCFGEVQLQPSLTP 120
Qy 121 LNRFTIWDVKAHKSIGLELOFSIPRLROIIGGESCPCDVYTHSISGRIDATVYRICTFSN 180
Db 121 LNRFTIWDVKAHKSIGLELOFSIPRLROIIGGESCPCDVYTHSISGRIDATVYRICTFSN 180
Qy 181 GTVSRIKQOEVEKALHLPMFHPNRVSGFSIANKSSIIRGLCTISVFEGESATLMSANY 240
Db 181 GTVSRIKQOEVEKALHLPMFHPNRVSGFSIANKSSIIRGLCTISVFEGESATLMSANY 240
Qy 241 PEGFEDELMTMOWFVPAHLRASVFLNFNLSNCRKERVEYIIPGSTTNPVEFKLBDK 300
Db 241 PEGFEDELMTMOWFVPAHLRASVFLNFNLSNCRKERVEYIIPGSTTNPVEFKLBDK 300
Qy 301 QPGNMGAFNLSLQCCDDAOSPGILRLOFVLYOHPONESNKIYVVDLSNBRASLTIE 360
Db 301 QPGNMGAFNLSLQCCDDAOSPGILRLOFVLYOHPONESNKIYVVDLSNBRASLTIE 360
Qy 361 PRPVAKSRKFVPGCFVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Db 361 PRPVAKSRKFVPGCFVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Qy 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVLPKXRLSLVLPKQLQOHTHEKPCNTSF 480
Db 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLVLPKXRLSLVLPKQLQOHTHEKPCNTSF 480
Qy 421 SYLVASAIIPSDLYFGSFCPGGSIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFLPY 540
Db 421 SYLVASAIIPSDLYFGSFCPGGSIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFLPY 540
Qy 541 FKEBGFVTPTDTSKAVLRTPNMDRGLPSLTSSVMNISVPRDOVACLTFFKERSGVVQ 600
Db 541 FKEBGFVTPTDTSKAVLRTPNMDRGLPSLTSSVMNISVPRDOVACLTFFKERSGVVQ 600
Qy 601 PRPVAKSRKFVPGCFVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Db 601 PRPVAKSRKFVPGCFVCLSESRCTSSNLTLSGSKHKISFLCDLTRLMMANVEKTIISCTDH 420
Qy 661 PRVTLVTLIAVGGVLLSALGLITCCVKKKKTKNKGPVAGIYXNNINTEMPROPK 720
Db 661 PRVTLVTLIAVGGVLLSALGLITCCVKKKKTKNKGPVAGIYXNNINTEMPROPK 720
Qy 721 KFOGRKONDSHVAVIEDTMYGHLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
Db 721 KFOGRKONDSHVAVIEDTMYGHLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
Qy 781 SRAPPAKLATEBPSPSPSESESEPTFSHPNNGDVSSKDTDIPILXTOEPMEPAE 836
Db 781 SRAPPAKLATEBPSPSPSESESEPTFSHPNNGDVSSKDTDIPILXTOEPMEPAE 836
```

RESULT 3

```
Q6NT71_HUMAN PRELIMINARY; PRT; 697 AA.
ID Q6NT71_HUMAN PRELIMINARY; PRT; 697 AA.
AC Q6NT71;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE CDCP1 protein (Fragment).
GN Name=CDCP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RN (1)
RX NUCLEOTIDE SEQUENCE.
RP TISUP-Kidney;
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.W., Feingold E.A., Grouse L.H., Berge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
```

RA Altechul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkin R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stadelton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huilyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Paley J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley A.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RP [2]
 RC NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RG NIH MGC Project;
 RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC069254; AAH69254.1; -, mRNA.
 FT NON TER 697 697
 SQ SEQUENCE 697 AA; 77703 MW; 9180BB3FEC3EB18 CRC64;

Query Match 82.9%; Score 3629; DB 2; Length 697;
 Best Local Similarity 99.7%; Pred. No. 3.4e-259;
 Matches 695; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGNCGVISALLGVLLIGARLPARGAEAFIALPRESNTIVLTKGTPTLLAKCYVI 60
 DB 1 MAGNCGVISALLGVLLIGARLPARGAEAFIALPRESNTIVLTKGTPTLLAKCYVI 60
 QY 61 SKRHITMSTKSGRIIVTFSCQSPENHFVIEIQKNDMSGPCPFGEVQLOPSTSLPT 120
 DB 61 SKRHITMSTKSGRIIVTFSCQSPENHFVIEIQKNDMSGPCPFGEVQLOPSTSLPT 120
 QY 121 LNRFTIMDVKAHKSIGLELQSPRIPLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDVKAHKSIGLELQSPRIPLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 QY 121 LNRFTIMDVKAHKSIGLELQSPRIPLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDVKAHKSIGLELQSPRIPLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 QY 181 GTVSRIRKQEGVKALHPWFHPRNVSGFSIANSSIRKLCIIESVFEGEGSATLMSANY 240
 DB 181 GTVSRIRKQEGVKALHPWFHPRNVSGFSIANSSIRKLCIIESVFEGEGSATLMSANY 240
 QY 241 PEGFPEDELMTQWVVPAPHLRASVSFLNPNLSNCRKEERYEYIIGSTTNPVEFKLEDK 300
 DB 241 PEGFPEDELMTQWVVPAPHLRASVSFLNPNLSNCRKEERYEYIIGSTTNPVEFKLEDK 300
 QY 301 QPGNAGNFNLSLQGCDDAOSPGLRLQFQVLVQHPONESNKIYVVDLSNERAMSLTIE 360
 DB 301 QPGNAGNFNLSLQGCDDAOSPGLRLQFQVLVQHPONESNKIYVVDLSNERAMSLTIE 360
 QY 361 PRPVQSRKFPVPGCFVCLIESRTCSSNLTITSGSKHKISFLCDLTRLMMNVEKTIISCTDH 420
 DB 361 PRPVQSRKFPVPGCFVCLIESRTCSSNLTITSGSKHKISFLCDLTRLMMNVEKTIISCTDH 420
 QY 421 RYCCRKSYSLQVPSIILHLPVELHDFSMKLLVPKRLSLVLPAPKQDQHPHEKCNISF 480
 DB 421 RYCCRKSYSLQVPSIILHLPVELHDFSMKLLVPKRLSLVLPAPKQDQHPHEKCNISF 480
 QY 481 SYLVASAIPODLYPSCFPGSIRKQIQVKONISVTLRTFAPSPFOEASROGLTYSFTLY 540
 DB 481 SYLVASAIPODLYPSCFPGSIRKQIQVKONISVTLRTFAPSPFOEASROGLTYSFTLY 540
 QY 541 FKESGVFTVPTKSKVYLRTPNMDRGLPSLTYSVSNISVPRDQVACLTFEKGSGVVCQ 600
 DB 541 FKESGVFTVPTKSKVYLRTPNMDRGLPSLTYSVSNISVPRDQVACLTFEKGSGVVCQ 600
 QY 601 TGRAFMIIQOORTAEIIFSLDEVDLPKPSFHHSFWNINSCSPSTSGKQDILFSLVLT 660
 DB 601 TGRAFMIIQOORTAEIIFSLDEVDLPKPSFHHSFWNINSCSPSTSGKQDILFSLVLT 660

DB 601 TGRAFMIIQOORTAEIIFSLDEVDLPKPSFHHSFWNINSCSPSTSGKQDILFSLVLT 660
 QY 661 PRVDLTITLLAANGGVLLLSAAGLITCCYKKKKKK 697
 DB 661 PRVDLTITLLAANGGVLLLSAAGLITCCYKKKKKK 697

RESULT 4
 Q6U9Y2_HUMAN
 ID Q6U9Y2_HUMAN PRELIMINARY; PRT; 691 AA.
 AC Q6U9Y2;
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE Membrane glycoprotein gp140 (Fragment).
 GN Name=CD81;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Epidermis;
 RX MEDLINE=96178080; PubMed=8647901; DOI=10.1083/jcb.132.4.727;
 RA Xia Y., Gil S.G., Carter W.G.;
 RT "Anchorage mediated by integrin alpha6beta4 to laminin 5 (epiligrin)
 RT regulates tyrosine phosphorylation of a membrane-associated 80-KD
 RT protein.";
 RL J. Cell Biol. 132:727-740 (1996).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Epidermis;
 RX PubMed=14739293; DOI=10.1074/jbc.M309678200;
 RA Brown T.A., Yang T.M., Zaitsevskaia T., Xia Y., Dunn C.A., Sigle R.O.,
 RA Knudsen B., Carter W.G.;
 RT "Adhesion or plaemin regulates tyrosine phosphorylation of a novel
 RT membrane glycoprotein p80/gp140/CUB domain-containing protein 1 in
 RT epithelia.";
 RL J. Biol. Chem. 279:14772-14783 (2004).
 DR EMBL; AY375452; AAR21289.1; -, mRNA.
 FT NON TER 691 691
 SQ SEQUENCE 691 AA; 76906 MW; 3B18AD27BDBF7338 CRC64;

Query Match 82.3%; Score 3603; DB 2; Length 691;
 Best Local Similarity 99.9%; Pred. No. 2.8e-257;
 Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGNCGVISALLGVLLIGARLPARGAEAFIALPRESNTIVLTKGTPTLLAKCYVI 60
 DB 1 MAGNCGVISALLGVLLIGARLPARGAEAFIALPRESNTIVLTKGTPTLLAKCYVI 60
 QY 61 SKRHITMSTKSGRIIVTFSCQSPENHFVIEIQKNDMSGPCPFGEVQLOPSTSLPT 120
 DB 61 SKRHITMSTKSGRIIVTFSCQSPENHFVIEIQKNDMSGPCPFGEVQLOPSTSLPT 120
 QY 121 LNRFTIMDVKAHKSIGLELQSPRIPLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 DB 121 LNRFTIMDVKAHKSIGLELQSPRIPLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 QY 181 GTVSRIRKQEGVKALHPWFHPRNVSGFSIANSSIRKLCIIESVFEGEGSATLMSANY 240
 DB 181 GTVSRIRKQEGVKALHPWFHPRNVSGFSIANSSIRKLCIIESVFEGEGSATLMSANY 240
 QY 241 PEGFPEDELMTQWVVPAPHLRASVSFLNPNLSNCRKEERYEYIIGSTTNPVEFKLEDK 300
 DB 241 PEGFPEDELMTQWVVPAPHLRASVSFLNPNLSNCRKEERYEYIIGSTTNPVEFKLEDK 300
 QY 301 QPGNAGNFNLSLQGCDDAOSPGLRLQFQVLVQHPONESNKIYVVDLSNERAMSLTIE 360
 DB 301 QPGNAGNFNLSLQGCDDAOSPGLRLQFQVLVQHPONESNKIYVVDLSNERAMSLTIE 360
 QY 361 PRPVQSRKFPVPGCFVCLIESRTCSSNLTITSGSKHKISFLCDLTRLMMNVEKTIISCTDH 420
 DB 361 PRPVQSRKFPVPGCFVCLIESRTCSSNLTITSGSKHKISFLCDLTRLMMNVEKTIISCTDH 420

|||||
Db PRPVQSKRFVPGCFVCSRTCSNLTSGSKHKISFLCDDLRLMNVNVEKTIISCDH 420
QY 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPCXRLSLVLPVPAQLOQHTHEKCNISF 480
Db 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPCXRLSLVLPVPAQLOQHTHEKCNISF 480
QY 481 SYLVASALPSODLYFGSFCPGSGIKQIQVKONISVTLRTFAPSFQOASRQGLTVSFIPY 540
Db 481 SYLVASALPSODLYFGSFCPGSGIKQIQVKONISVTLRTFAPSFQOASRQGLTVSFIPY 540
QY 541 FKEEGVFVTPTKSKVYLRTPNMDRGPLSLTSVSNISVPRDQVACLTFFKERSGVQC 600
Db 541 FKEEGVFVTPTKSKVYLRTPNMDRGPLSLTSVSNISVPRDQVACLTFFKERSGVQC 600
QY 601 TORAFMIIOEORTRAEEIFSLDEVLPRKSEFHHSFWNINSCSPSGKQDLSFVTLT 660
Db 601 TORAFMIIOEORTRAEEIFSLDEVLPRKSEFHHSFWNINSCSPSGKQDLSFVTLT 660
QY 661 PRTVDLTVLLIAAVGGVLLLSALGLIICV 691
Db 661 PRTVDLTVLLIAAVGGVLLLSALGLIICV 691

RESULT 5
OSU462 MOUSE PRELIMINARY; PRT; 833 AA.
ID OSU462
AC OSU462
DT 01-FEB-2005 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
DE 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DB CUB domain-containing protein 1.
GN Name=Cdcp1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schletter G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustun T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.T., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Maitra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RG NIH MGC Project;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC085253; AAH85253.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR000859; CUB.
SQ SEQUENCE 833 AA; 92725 MW; 1F1423DD595C6B27 CR64;

Query Match 81.7%; Score 3574.5; DB 2; Length 833;

Best Local Similarity 80.9%; Pred. No. 4.7e-255;
Matches 676; Conservative 60; Mismatches 97; Indels 3; Gaps 3;

QY 1 MAGLNCGSIALGVLILGAAARLPAGAFAPALPRESNITVLTGLSTPTLAKCYIVI 60
Db 1 MAHSACGFSVALLGALLGTRARLLNGTEASEIALPQRSGVVISILGNPALPVKCYIWM 60
QY 61 SKRHITMISISGERIVETFCQSPENHFVIEIQNIQCMGAPCPFGVQLOPSTSLPT 120
Db 61 SRQHTTEILIRPEKKSFTFCSNPEKHFVLKLEINIDCMGCPFGVHLOPSSLPI 120
QY 121 LNRFTIMDKAKHSIGLELQFSIPRLRQIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTIMDKAKHSIGLELQFSIPRLRQIGGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSIRLCIIEVFEGBSATLMSANY 240
Db 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANRSIRLCIIEVFEGBSATLMSANY 240
QY 241 PEGFPEDELMTWQFVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTNPVEFKLBDK 300
Db 241 PEGFPEDELMTWQFVPAHLRASVSFLNPNLSNCRKEERYEYIPGSTTNPVEFKLBDK 300
QY 301 QPGNMAGNFNLSLQCCDDAQSFGILRLQFVLVQHPQNESNKIYVVDLSNERAMSLTIE 360
Db 301 QPGNMAGNFNLSLQCCDDAQSFGILRLQFVLVQHPQNESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSKRFVPGCFVCSRTCSNLTSGSKHKISFLCDDLRLMNVNVEKTIISCDH 420
Db 361 PRPVQSKRFVPGCFVCSRTCSNLTSGSKHKISFLCDDLRLMNVNVEKTIISCDH 420
QY 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPCXRLSLVLPVPAQLOQHTHEKCNISF 480
Db 421 RYCQKRSYSLQVPSDILHLPELHDFSWKLLVPCXRLSLVLPVPAQLOQHTHEKCNISF 480
QY 481 SYLVASALPSODLYFGSFCPGSGIKQIQVKONISVTLRTFAPSFQOASRQGLTVSFIPY 540
Db 481 SYLVASALPSODLYFGSFCPGSGIKQIQVKONISVTLRTFAPSFQOASRQGLTVSFIPY 540
QY 541 FKEEGVFVTPTKSKVYLRTPNMDRGPLSLTSVSNISVPRDQVACLTFFKERSGVQC 600
Db 541 FKEEGVFVTPTKSKVYLRTPNMDRGPLSLTSVSNISVPRDQVACLTFFKERSGVQC 600
QY 601 TORAFMIIOEORTRAEEIFSLDEVLPRKSEFHHSFWNINSCSPSGKQDLSFVTLT 660
Db 601 TORAFMIIOEORTRAEEIFSLDEVLPRKSEFHHSFWNINSCSPSGKQDLSFVTLT 660
QY 661 PRTVDLTVLLIAAVGGVLLLSALGLIICVKKKKKTKNKGPVAVGIYXNINTEMPRQPK 720
Db 661 PRTVDLTVLLIAAVGGVLLLSALGLIICVKKKKKTKNKGPVAVGIYXNINTEMPRQPK 720
QY 721 KFOQRKNDSHVAVVIEDTWYVGHLLDSSGSLQPRVDPVYRPRQGMGCPSPPTIC 780
Db 721 KFOQRKNDSHVAVVIEDTWYVGHLLDSSGSLQPRVDPVYRPRQGMGCPSPPTIC 780
QY 781 SRAPFAKLATEPRSPSPSESESEPTPSHNNNGVSKODPIPLTXQEPMEPAE 836
Db 781 SRAPFAKLATEPRSPSPSESESEPTPSHNNNGVSKODPIPLTXQEPMEPAE 836
QY 836 SRAPFAKLATEPRSPSPSESESEPTPSHNNNGVSKODPIPLTXQEPMEPAE 836
Db 836 SRAPFAKLATEPRSPSPSESESEPTPSHNNNGVSKODPIPLTXQEPMEPAE 836
RESULT 6
OS098C2 HUMAN PRELIMINARY; PRT; 649 AA.
ID OS098C2
AC OS098C2
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DB Hypothetical protein FL13772.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Homo.
OX NCBI_TaxId=9606;


```

RN      [1]
RC      NUCLEOTIDE SEQUENCE.
RC      TISSUE=Placenta; DOI=10.1038/ng1285;
EX      PubMed:14702039;
RA      Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA      Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA      Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA      Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA      Negahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA      Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA      Sugawara M., Takahashi K., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA      Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA      Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA      Fujimori K., Tanai H., Kimata M., Watanabe M., Hirooka S., Chiba Y.,
RA      Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotta T.,
RA      Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA      Nomura Y., Matsushino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA      Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA      Moriya S., Momiyama H., Satoh N., Takami S., Teraehima Y., Suzuki O.,
RA      Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA      Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA      Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA      Fujimori Y., Komiyama M., Tashiro K., Tanigami A., Fujiwara T.,
RA      Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Omori Y.,
RA      Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA      Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA      Matsunura R., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA      Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA      Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA      Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA      Nakai K., Yada T., Nakamura Y., Ohara O., Isegai T., Sugano S.,
RT      "Complete sequencing and characterization of 21,243 full-length human
RT      cDNAs."
RL      Nat. Genet. 36:40-45(2004).
DR      EMBL: AK023834; BAB14695.1; -; mRNA.
DR      GO: GO:0016021; C: integral to membrane; IEA.
DR      InterPro: IPR000859; CDB.
SQ      SEQUENCE 649 AA; 72579 MW; 0F4404AD31D60E9A CRC64;

Query Match      77.6%; Score 3395; DB 2; Length 649;
Best Local Similarity 99.2%; Pred. No. 6,1e-242;
Matches 644; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

DB      |||||
DB      IOEORTRAEEIFSLDEDEVLPKPSFHHHSFWNINSCSPTSKQDLPLFSVLTLPRTVDT 480
QY      421
QY      VILAAVGGVLLSALGLIICVKKKKKKTKNGPAGVIYNNXINTENMPROPKKFOGKRK 727
DB      468
DB      VILIAAVGGVLLSALGLIICVKKKKKKTKNGPAGVIYNNXINTENMPROPKKFOGKRK 540
QY      481
QY      728
QY      DNDSHVAVIEDTMYGHLLODSSGSLQPEVDYTRPPQGTGVCPPSPPTICSAFPAK 787
DB      541
DB      DNDSHVAVIEDTMYGHLLODSSGSLQPEVDYTRPPQGTGVCPPSPPTICSAFPAK 600
QY      788
QY      LATEPPRPSPSESEPEPTTSHPNNGVSSKDDTDIPLLTQEPMEPAE 836
DB      601
DB      LATEPPRPSPSESEPEPTTSHPNNGVSSKDDTDIPLLTQEPMEPAE 649

RESULT 7
ID      Q9H676 HUMAN PRELIMINARY; PRT; 392 AA.
AC      Q9H676;
DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT      01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE      Hypothetical protein FLJ22534.
OS      Homo sapiens (Human)
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC      Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA      Okitani R., Ota T., Suzuki Y., Ohyashi M., Nishi T., Shibahara T.,
RA      Tanaka T., Nakamura Y., Isegai T., Sugano S.,
RL      Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AK026187; BAB15388.1; -; mRNA.
FT      NON TER
SQ      SEQUENCE 392 AA; 44030 MW; 3E21CBB2703E52F3 CRC64;

Query Match      46.6%; Score 2041; DB 2; Length 392;
Best Local Similarity 99.5%; Pred. No. 3,6e-142;
Matches 390; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

RESULT 8

OX NCBI_TaxID=9606;
 RP [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RG NIH MGC Project;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
 RA Clark H.F., Gurney A.L., Adaya E., Baker K., Baldwin D.T., Brush J.,
 RA Chen J., Chow B., Chui C., Crowley C., Curral B., Denel B., Dowd P.,
 RA Batson D., Foster J.S., Grimaldi C., Gu Q., Hase P.E., Heidens S.,
 RA Huang A., Kim H.S., Klimowicz L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M.R., Robble B., Sanchez C., Schoenfeld J.,
 RA Seehaght S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R.L., Watnabe C., Wleand D., Woods K., Xie M.-H.,
 RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
 RA Wood W.I., Godowski P.J., Gray A.M.;
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genom. Res. 13:2265-2270 (2003).
 RL EMBL; BC021099; AAH21099.1; -; mRNA.
 DR EMBL; AY358779; AA089139.1; -; mRNA.
 DR Ensembl; ENSG00000163814; Homo sapiens.
 SO SEQUENCE 343 AA; 37818 MW; 3E9E13379DD94D1B CRC64;
 Query Match 40.6%; Score 1779; DB 2; Length 343;
 Best Local Similarity 99.4%; Pred. No. 6; 9e-123;
 Matches 341; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MAGLNCGVSIALGLVLLGAARLPFGAFAFIALPRESNTITLKLGPPTLAKCYIV 60
 DB 1 MAGLNCGVSIALGLVLLGAARLPFGAFAFIALPRESNTITLKLGPPTLAKCYIV 60
 OY 61 SKRHTMTSISGERIVTFPSCQSPENHFVIEIQKNICMSPCFGEVQLOPSTSLPT 120
 DB 61 SKRHTMTSISGERIVTFPSCQSPENHFVIEIQKNICMSPCFGEVQLOPSTSLPT 120
 OY 121 LNRTFTMDVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 DB 121 LNRTFTMDVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 OY 121 LNRTFTMDVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 DB 121 LNRTFTMDVKAHKSIGLELOFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
 OY 181 GTVSRITKQEGVKALHPWPHPRVSGPSIANRSIKRLCTISVEFEGEGSATIMSANY 240
 DB 181 GTVSRITKQEGVKALHPWPHPRVSGPSIANRSIKRLCTISVEFEGEGSATIMSANY 240
 OY 241 PEGFPEDELMTWQFVPAHLRASVSFLNFNLSNCRKEERVEYYPGSTTNEPVFKLBDX 300
 DB 241 PEGFPEDELMTWQFVPAHLRASVSFLNFNLSNCRKEERVEYYPGSTTNEPVFKLBDX 300

DB 241 PEGFPEDELMTWQFVPAHLRASVSFLNFNLSNCRKEERVEYYPGSTTNEPVFKLBDX 300
 OY 301 QPNNMAGNPNLSLOQCDDAOSPGILRLQFOVLVQHPONESNK 343
 DB 301 QPNNMAGNPNLSLOQCDDAOSPGILRLQFOVLVQHPONESSE 343
 RESULT 10
 ID Q921M9_MOUSE PRELIMINARY; PRT; 384 AA.
 AC Q921M9;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Cdcp1 protein (Fragment).
 GN Name=E030027H19RLK; Synonyms=Cdcp1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RG NIH MGC Project;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC011340; AAH11340.1; -; mRNA.
 DR EMBL; ENSMUSG000000035498; Mus musculus.
 DR MGI; MGI:2442010; Cdcp1.
 DR MGI; MGI:2442010; E030027H19RLK.
 FT NON_TER 1
 SO SEQUENCE 384 AA; 42406 MW; 4B5F17D30FE16080 CRC64;
 Query Match 36.1%; Score 1579.5; DB 2; Length 384;
 Best Local Similarity 77.3%; Pred. No. 4; 6e-108;
 Matches 299; Conservative 34; Mismatches 51; Indels 3; Gaps 3;
 OY 450 LLVPKDRISLVLPAPQKIQOHTHEKPCNTSSYLVASAIPSQDLYFGSFCGSGSIKQIOV 509
 DB 1 LLVPKDRISLVLPAPQKIQOHTHEKPCNTSSYLVASAIPSQDLYFGSFCGSGSIKQIOV 509
 OY 510 KQNTSVTLRTAPSPFOEASRQGLTVSFIPEKKEGVFTVPPDRSKVYLTATPMMDGLP 569
 DB 510 KQNTSVTLRTAPSPFOEASRQGLTVSFIPEKKEGVFTVPPDRSKVYLTATPMMDGLP 569
 OY 61 KQNTSVTLRTAPSPFOEASRQGLTVSFIPEKKEGVFTVPPDRSKVYLTATPMMDGLP 120
 DB 61 KQNTSVTLRTAPSPFOEASRQGLTVSFIPEKKEGVFTVPPDRSKVYLTATPMMDGLP 120
 OY 570 SLTVSNVNIIVPRDOVACLTFPKERSGVVCGTGRAFMIIORATRAEIRFSLDQVLPKP 629
 DB 570 SLTVSNVNIIVPRDOVACLTFPKERSGVVCGTGRAFMIIORATRAEIRFSLDQVLPKP 629
 OY 121 ALSSVSNVNIIVPRDOVACLTVLTKERSGLACGSAFMIIQDQSRABEIRFSLDQVLPKP 180
 DB 121 ALSSVSNVNIIVPRDOVACLTVLTKERSGLACGSAFMIIQDQSRABEIRFSLDQVLPKP 180

```

Oy 630 SFHHHSFWWNISNCSPTSGKOLDLFSVTLTPRTVDLTVILIAVGGVLLLSALGLIIC 689
Db 181 SFHHHSFWWNISNCSPTSGKOLDLFSVTLTPRTVDLTVILIAVGGVLLLSALGLIIC 239
Oy 690 CVKKKKKKTKNGKPAVGIYNNXINTEMPRPKKFKGKRDNDSHVAVIEDTWYGHLLD 749
Db 240 FV-KKKKKKVDGPAVGIYNNXINTEMPRPKKFKGKRDNDSHVAVIEDTWYGHLLD 297
Oy 750 SSGSLQLOEVDVTPRFQGTGVCPPSPPTICSRATPAKLATEPRRSPRESSEPTYES 809
Db 238 SSGSLQLOEVDVTPRFQGTGVCPPSPPTICSRATPAKLATEPRRSPRESSEPTYES 357
Oy 810 HPNNGDVSSKDTDIPDLXTOEBEPAPAE 836
Db 358 HPNNGEIGVRETDIPDLXTOEBEPAPAE 384

RESULT 11
OAT223 TETNG PRELIMINARY; PRT; 867 AA.
ID OAT223 TETNG PRELIMINARY; PRT; 867 AA.
AC OAT223 TETNG PRELIMINARY; PRT; 867 AA.
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Chromosome undetermined SCAF10362, whole genome shotgun sequence.
GN ORFNames=GSTENG0000853001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxId=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Anry J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozout-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dastiva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vachele B.,
RA Bismont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
RA Crnaud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Landet V., Schachter V., Querier F., Saurin W., Scarpelli C.,
RA Winkler P., Landet E.S., Weissenbach J., Roest Crolious H.,
RA RT the early vertebrate proto-karyotype."
RA RT Nature 431:946-957 (2004).
RN [12]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CAAB01010362; CAP3059.1; -: Genomic DNA.
SQ SEQUENCE 867 AA; 96120 MM; 1177A6DEBABA88A CRC64;

Query Match 18.9%; Score 828; DB 2; Length 867;
Best Local Similarity 30.0%; Pred. No. 5.4e-52;
Matches 243; Conservative 135; Mismatches 311; Indels 120; Gaps 32;

```

```

Db 220 VLYKALMTLQVGHRELDVVDFTLSNGEPTMTAVKSLPRGVSTGFFAANYPNDEPD 279
Oy 247 DELMTWQVNVFAHLBASVFLNPNLSNCRKEEVEYIIPGSTNPEVFKLEDKQPGMA 306
Db 280 KOQVEWBDVTVGMENYVHFQAHFAPBCLNGKEVEEYHKKQSKVTR--LALTDPOBEHQ 337
Oy 307 GNPNLSLQGCQDQASPGILRLQFOVLVQ--HPONESNKIYVVDLSNERAMSLTIEBRP 363
Db 338 GDFKAVLKNKGTNTNLTQGLTLDYKVSVMRSGHP-----VLCTVDLTGKQVSLDQ---- 387
Oy 364 VKQSKRFLVPGFVCLIESFTCSNLTLSGSKHKSIFLCLDDLRLMNVKXITISCDHXYC 423
Db 368 LKVGSD--PYCEMSVNSK-IEKINVAAGTILASISFLDCCPQDVALTATKTIACHTLSC 444
Oy 424 QKSYSLQVPSDILHLPELHDFSWKLVLPKDRSLVLPVPAQKLOQHTHEKPCNTSFSYL 483
Db 445 --PSSLTVPBLSCLPLPHSHFTMHMMIIPDSTVDLMSPGSLHQSIPQGECSLSLH 502
Oy 484 VASAIPQDLYFGSFPCGSIKQIQVKNISVTLRTFAPSFXOASROGLTVSFLPYRKE 543
Db 503 VAE--SDGLSLGDFCFNGALQKIOAHNISTAR--VPDF-KKSRGPFLLASFQGEIOD 556
Oy 544 EGFTVTPDYTSKYVLRTPNMDRGLPSLTYSWNISVPRDOVALCTFKKESGVVC--Q 600
Db 557 TFIYRISPE-EPQTLATPMPQGMKPSSTVSWVTLPSQYEAADRFFNV-SQPKDMSH 614
Oy 601 TGRAMIIQEQRTAREEIPSLDED-----VLPRKSFHHHSFWWNISNCSPTSGKOLDL 653
Db 615 TSIKVMTLQGE--DELSRREDPKYDMLVPE-----SFIYMNKCLPQEG-QRKA 663
Oy 654 LFSVTLTPRTVDLTVILIAVGGVLLLSALGLIICVKKKK--KTKGPAVGIYNNXNI 711
Db 664 MAKIVLQKSNLPLIIL--GVAQAFLLILLIILAVVCVKKKKSKKANKNSIYWKAKSI 721
Oy 712 NTEMPOQKFKQ-KCRKNDSHVAVIEDTWYGHLLDSS-----GSFLQPEVDYTRP 764
Db 722 ----FRDEIRFKRSRDNNSHIYDSIDETWYGHLLPDSYSTLSHSGMYDSYOT 776
Oy 765 FQG-TMGVCPSPPTICSRAPPAKLATEP-----PPRS--PPES----- 801
Db 777 FSGPTDGKLP-----VIEPDHDEVEQGTFLDPKRSFMPRPRTPIDR 821
Oy 802 -----ESPEYTFSHPNNDVSS 818
Db 822 QDSLGFQDSRMVNDNLYTFK--STGDWNT 848

RESULT 12
OAT223 TETNG PRELIMINARY; PRT; 825 AA.
ID OAT223 TETNG PRELIMINARY; PRT; 825 AA.
AC OAT223 TETNG PRELIMINARY; PRT; 825 AA.
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Chromosome undetermined SCAF7985, whole genome shotgun sequence.
GN ORFNames=GSTENG0000554001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxId=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Anry J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozout-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dastiva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vachele B.,
RA Bismont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
RA Crnaud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,

```


DT 10-MAY-2005 (TrEMBLrel. 30, last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, last annotation update)
 DE Hypothetical LOC541520.
 GN Name=zgc:113071;
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NC NCBI_Taxid=7955;
 OK [1]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=Singapore local strain; TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheifter C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussidi T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J.,
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield J.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=Singapore local strain; TISSUE=Embryo;
 RG NIH MGC Project;
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC091838; AAI91838.1; -, mRNA.
 KW Hypothetical protein.
 SQ SEQUENCE 793 AA; 88434 MW; C941FC39F0F9F3A9 CRC64;
 Query Match 13.9%; Score 610; DB 2; Length 793;
 Best Local Similarity 25.7%; Pred. No. 6,1e-36;
 Matches 215; Conservative 153; Mismatches 372; Indels 98; Gaps 30;
 QY 9 STALLGVLILGAAFLPRGAFAEIALPRESNITVLIKIGTPT-LIAKPC----- 56
 DB 10 SLVILVFSLSVLELTGAE-LINVAV--EQGTIILNTSSSTSTQCKVMRYANQLCOA 66
 QY 57 YIVISKRTITMISISGERIVTFPSCQSPENHFVIEIOKNIDCMSPGCFGEVQLDPT- 115
 DB 67 YVVL--RSPTLDD-----FSCSQPEKVFIVQIDPT-----GTMTLEPKCH 105
 QY 116 SLPLTLNRTFIDVKAHKSIGLEQSIPLRLQIGSGSCPDGVTISGRIDATVVRIG 175
 DB 106 SAPIKRNRTFNKRLKLVSKSLQIANSSSGIKQIREDSCPDNRKRIIA-----VENISIG 161
 QY 176 TFCNSGTVARIKMQEGVAKALHPWHPRNVSFGSIANSSIKRLCIISVEGEGSATL 235
 DB 162 TFCQNGVIRQIDVRNAGKLSAEVTGGQSISTKVIIVSVGLMNSLVNIHAIIVPEKSS--I 219
 QY 236 MEANYPEGFEDELMTWQVVPRAHLASVSLFNFNISNCKEKE--RVEYIIPGSTTNE 293
 DB 220 QDEITPNTTPENAETWTFMFSMPAAVYADRIILKYTLPTCLISBSHPIPMKTYWG--RDAL 277
 QY 294 VKLEEDKQGNMAFNLSLOGDDQ--AOSPGLRLQFOFV-LVQHPQNSNKKLYVVDLSN 351
 DB 278 VKRMENQSVBEGANNISIKCKMSTVSLNRLMHPQISAIKRRKGCEG----DLFD 333
 QY 352 ERAMSLTIBPRPKQSRKVPGGCVLESRTGSSNLTLSGSKHKISPLCDLITRLMMNV 411
 DB 334 KQVLKIQV---MKDPKSV--CVLKLDS-VIMDTVTIASGNFIINTFDCKNDELRLTV 386

QY 412 EKTISTDHRYSQQRKSYSLQVPSDILHPVELHDPFSKTLVPKDLSTVLVPAQLQOHT 471
 DB 387 IQTECKWKNCSTSTLFLHNYEQQCIIPGVLMKMTWYIQGPVNAVLSQSTDLARCL 446
 QY 472 HEKPCNTSPSYLVASAPSDLYFGSPFGGSIKIQVKN-ISVTLTTPAFSFXQASR 530
 DB 447 PEDNCN-STILLNVSQVWNPPTTVQFCPKGSVQKIQIRESKIATVASTSPSDRSVAK 505
 QY 531 QGLTVSFTPYKEBEVFTVTDITKSKVLTIRPNMGRGLPSLTSVSNMNSVPRDQVACLTF 590
 DB 506 PILSYSTFOGISSENIYFVTPMDPTTLATPGWPSAKASSTVSWINMEPQYTSVLT- 564
 QY 591 FKERSGVACQGRARFMIIOEORTAEIIFSLDEVLPRPSFHHSFWMNISNC-SPTSGK 649
 DB 565 FRNVQPKCKQVHTNIAVQITRSQ-VLYTTKDKMDLILBEFYLMNTCKSPITGA- 622
 QY 650 QLDLFSVTLTPRTVDTLVIIIAVGGVLLISALGLIICVKKKKKKTKNKPAGVGINX 709
 DB 623 -FRAMWETVQSNNTKLGIIIGVY--GVVLAIAVAVIWFVLRKKR-NKAPVSVYNP 678
 QY 710 NINTMPROPKFKQGRKNDNSHYAVIEDTMYGHILQDSSGSLQDPEVDITRFFQGTM 769
 DB 679 SEHAPLPGIHGIPKQGEDEENHTYAVYIDTLVYSHLND-----DADKKQHYEDTS 730
 QY 770 GVCPSPP-----TICSRAPTAKL-----ATEEPPRSPSESESEBYTF 808
 DB 731 GPHLAKPPLPDRPDNDVSHALDKLLPNELYGYVGQSTSGPFRTRAGKS-SEHTY 787
 RESULT 15
 001654 HALRO
 ID 001654 HALRO PRELIMINARY; PRT: 746 AA.
 AC 001654;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE AEMASPa.
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyuridae; Halocynthia.
 NC NCBI_Taxid=7729;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Hepatopancreas;
 RX MEDLINE=97322374; PubMed=9177219; DOI=10.1073/pnas.94.12.6340;
 RA Uj X., Azumi K., Sasaki M., Nonaka M.,
 RT "Ancient origin of the complement lectin pathway revealed by molecular
 RT cloning of mannan binding protein-associated serine protease from a
 RT urochordata, the Japanese Ascidian, Halocynthia roretzi.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:6340-6345(1997).
 DR EMBL, D88204; BAA19762.1; -, mRNA.
 DR HSP; G9JUS8; INT0.
 DR GO; GO:0005576; C:extracellular region; IEA.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0004263; F:cyclo-oxygenase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006956; P:complement activation; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000152; Asx_hydroxyl_5.
 DR InterPro; IPR001881; EGF Ca.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR001254; Peptidase_S1_96.
 DR InterPro; IPR011361; Pept_S1_Comp_Act.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00089; Sushi; 1.
 DR Pfam; PF00089; Trypsin; 1.
 DR PIRSF; PIRSF001155; C1r_C1s_MASP; 1.
 DR PIRSF; PIRSF001155; C1r_C1s_MASP; 1.
 DR SMART; SM00032; CCP; 2.
 DR SMART; SM00042; CUB; 2.

DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00010; ASK_HYDROXYL; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50923; SUSH1; 2.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
SQ SEQUENCE 746 AA; 84127 MW; 32B075317CF173E1 CRC64;

Query Match 3.2%; Score 141; DB 2; Length 746;
Best Local Similarity 19.3%; Pred. No. 0.26;

Matches 102; Conservative 73; Mismatches 178; Indels 176; Gaps 25;

QY 220 LCITISVEEGEGSATLMSA-----NYEGFPEDELTWQFVVPAAHLRASVSFLNENTL 271
DB 11 LLYCNSVF-----SAILLTAHFQNFSSPNYPSPDNSNLTNIRVQHGYRMSIRPSTFDL 66
QY 272 -----SNCRKERVEEYIRGSTTNEVEFKLEDKOPGNMAGNPNLSLQCCDODADOSP 323
DB 67 EDSYEDGIGSC-----YDYVEITESNKTVAK-----FCGNVQL-----FPTDAPNP 108
QY 324 GILRLQFQVLVQHPONESNKIYVVDLS-----NERAMSLTIEPPRVKOSRK----- 369
DB 109 -----SKFITSQNEVRVTFVSDYSLSGFOAHYAQIDINECELMETKRTIIEDMD 161
QY 370 -----FVPGCFVCLIESRTCSNLTLSGSRHKISFLCDDLTRLMNVKTIISCTD 419
DB 162 ELVVCASHYCRNVPGSYYC-----GCRPKFTLDNRHTCVASFCEN-----QVLTDDN 208
QY 420 HRYCCOKSKYSLOVPDILHLPEVLHDFSWKLLVPRD-RLSILVLPAAQKLQCHTHEKPCNT 478
DB 209 SGHISPEP-----ELVAKLTDCSWTIQLREGLSVNLIFERAFGIEHEEE--GC 257
QY 479 SFSYLVASAIPESODLYFGSFCPCG-----SIKOIVKONISYTLRTFAPSFXQ---- 526
DB 258 TYDRLAEVLHKSTTDLYCGNQAPGNGTVMPMNTNLVQLKHTDLSVEKKGFSVRYTSTRIK 317
QY 527 -----EASRQGLTVSFI--PYFKEGVFTVTPDTKSKVYLRTPNMDRGLPSLTSV---S 575
DB 318 CLHALHDPKNG-SLSFSHSRSYHEFEDVATFSCDRGFDLI-----GVPRKCLSDGS 368
QY 576 WNISVPRDQV-AC-----LTFEKRSGVVCQ-----TGRAFMII 608
DB 369 WSHSAPICQIKSCGVPOFLDLPNSHIVEYENSKTYSBVLDTVCNQWYGMISGASKMIC 428
QY 609 QEORTRAEEIFSLDEDLVLPKPSFHHSFPWNISN-----CSPTSGKQJ 651
DB 429 ENSKIWTE-----HGGIVAINNFNNKPVCKPICGKTL 460

Search completed: May 4, 2006, 22:15:05
Job time : 125 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 4, 2006, 21:01:05 ; Search time 49 Seconds

(without alignments)
1410.548 Million cell updates/sec

Title: US-10-781-564-1

Perfect score: 4392
Sequence: 1 MAGUNCGVSTLLGVLLGA.....SSKDTDIPLSTQPMPEAE 836

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%
Maximum Match 100%

Listing first 1000 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

No matches found

Search completed: May 4, 2006, 21:02:36

Job time : 49 secs

THIS PAGE LEFT BLANK